

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2005, 19:08:10 ; Search time 42 Seconds
(without alignments)
1459.288 Million cell updates/sec

Title: US-10-623-429-9

Perfect score: 3385
Sequence: 1 MASAEMRERLEAPLDRAPV.....VNVDTARAADLFVSMWMSR 637

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR.79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2333	68.9	635	1 WMBE66	capsid protein - h
2	1262	37.3	306	2 A54862	HSV-1 proteinase -
3	1079.5	31.9	646	1 WZBSC8	68.6K capsid prote
4	1056	31.2	647	2 T42579	capsid protein 35
5	943	27.9	522	2 S52216	viral proteinase -
6	824.5	24.4	605	1 WZB333	gene 33 protein -
7	401	11.8	586	1 A43675	polyprotein - equi
8	383.5	11.3	643	2 S55610	capsid protein p40
9	372.5	11.0	605	1 QQB33R	BVPF2 (BC-RF3) pro
10	331	9.8	459	2 T42930	hypothetical prote
11	322	9.5	590	1 WMBSCB	64K capsid assembl
12	317.5	9.4	708	1 QQB388	UL80 protein - hum
13	301.5	8.9	475	1 WZBEM4	gene 17 protein -
14	280	8.3	839	2 T04859	extensin homolog F
15	278	8.2	524	2 T03112	minor capsid scaff
16	278	8.2	697	2 T44013	fused proteinase-c
17	272.5	8.1	528	2 T44013	proteinase / in-fr
18	270.5	8.0	528	2 T44200	probable proteinase
19	269	7.9	620	2 S06733	hydroxyproline-ric
20	264	7.8	2715	2 T13049	eyelid - fruit fly
21	262.5	7.8	519	2 S55470	proteinase - human
22	258.5	7.6	760	2 T06291	extensin homolog T
23	254.5	7.5	574	2 T43556	Wiskott-Aldrich sy
24	254.5	7.5	574	2 T38819	Wiskott-Aldrich sy
25	253.5	7.5	1184	2 G01763	atrophin-1 - human
26	251	7.4	464	2 S22597	extensin - Volvox
27	250	7.4	1494	2 T14355	protein-tyrosine-p
28	247.5	7.3	1184	2 S50832	atrophin-1 - human
29	246.5	7.3	1188	2 S49915	extensin-like prot

30	245.5	7.3	727	2 C84534	hypothetical prote
31	244	7.2	801	2 T29018	hypothetical prote
32	243.5	7.2	1541	2 T02831	AAA protein L4171.
33	242.5	7.2	1585	2 T31611	hypothetical prote
34	242	7.1	895	2 C86371	99.7K hypothetical
35	240	7.1	660	1 QQB33	BHLFI protein - hu
36	239	7.1	322	2 S25299	extensin precursor
37	238	7.0	786	2 T01456	extensin homolog F
38	237.5	7.0	744	2 E86255	hypothetical prote
39	237	7.0	512	2 T41955	probable proteinase
40	236.5	7.0	1201	2 G86441	unknown protein [i
41	235	6.9	278	2 S20790	extensin - almond
42	234.5	6.9	1952	2 T48814	hypothetical prote
43	233.5	6.9	317	2 A28996	proline-rich prot
44	233.5	6.9	733	2 T47618	extensin-like prot
45	233.5	6.9	3164	1 WMBE66	UL36 protein - hum

ALIGNMENTS

RESULT 1

WMBE66

capsid protein - human herpesvirus 1 (strain 17)

C;Species: human herpesvirus 1

C;Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 09-Jul-2004

C;Accession: H30084

R;McGeoch, D.J.; Dalrymple, M.A.; Davison, A.J.; Dolan, A.; Frame, M.C.; McNab, D.; Perr

J. Gen. Virol. 69, 1531-1574, 1988

A;Title: The complete DNA sequence of the long unique region in the genome of herpes sim

A;Reference number: A30083; MUID:88274327; PMID:2839594

A;Accession: H30084

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-635 <WCG>

A;Cross-references: UNIPROT:P10210; GB:X14112; NID:G1944536; PIDN:CAA32318.1; PID:G59526

C;Genetics:

C;Superfamily: varicella-zoster virus gene 33 protein

C;Gene: UL26

C;Keywords: capsid protein

Query Match 68.9%; Score 2333; DB 1; Length 635;
Best Local Similarity 74.0%; Pred. No. 6.6e-114;
Matches 477; Conservative 30; Mismatches 120; Indels 18; Gaps 7;

Qy	1	MASAEMRERLEAPLDRAPVIYVAGFLALYDSGDELALDPDVTVAALPENPLINVD	60
Db	1	MAADAPGDRMEELPDRAPIYVAGFLALYDSGDELALDPDVTVAALPDNPLINVD	60
Qy	61	HRARCEVGRVLAIVNDPRGPFVGLIACVQLERVLETAASAAIFERRGPALSREERLLYL	120
Db	61	HRAGCEVGRVLAIVDDPRGPFVGLIACVQLERVLETAASAAIFERRGPALSREERLLYL	120
Qy	121	ITNYLPSVSLSTRGDEVPDRTLFAHVALCAIGRLGTIVTYTSLDAAIAFFRHLP	180
Db	121	ITNYLPSVSLATKRLGAEHPDRTLFAHVALCAIGRLGTIVTYTGLDAAIAFFRHLP	180
Qy	181	ATREGVRRERAAELALAGRTWAPGVEALTHLLSTAVNNMLDRWSLVAERROAGIA	240
Db	181	ASREGARRLAAEAELALSGRTWAPGVEALTHLLSTAVNNMLDRWSLVAERROAGIA	240
Qy	241	GHTYLAQSEKFKIWGAESAAPAPRGYKTGAPGAMDTSPAASVPAPO----VAVRAROVAS	296
Db	241	GHTYLAQSEKFKWGAESAAPAPRGYKNGAPGPESTDIPPGSIAAAPQGDRCFPIVRQGV	300
Qy	297	SSSSSFFPAPADNMNVSASGAPPPPGSGSYLWIPASHYNQLVTGOSAPRHPPLTACGL	356
Db	301	S-----PVLPMNPVPTSGTPAPAPPGSGSYLWIPASHYNQLVAGHAAPQPQPHSAFG	354
Qy	357	P-AAGTVAVCHPGAGSPHYPPPPHPYPCMLFAGSPLEAQIAALVGAIAAARQAGLP	415
Db	355	PAAAGVAVYGPAGLGSQHYPPHVAHQYPGVLFGSPLEAQIAALVGAIAAARQAGQP	414

Qy	416	AAAGDGHGIRGSAKRHRHEVEQDEYDCGRDEPD	FPYYPGEARPEPPVDSRRAARQASG	475
Db	415	-AAGDPGVGRSGKRRYRAGPSESYCDQDEPDADY	FPYYPGEARGARGVDSRRAARHSPG	473
Qy	476	PHEITITLGVAVTSLQOELAHMRARTAPYGPVP	GVGYPHHPHADTE-TPAQPPRYPAKA	534
Db	474	TNETITLGVAVTSLQOELAHMRATSPYGMVTP	VAHYRYPQVGEPEFTTHPALCPPEA	533
Qy	535	VYLPPPHIAPPGLPSGA--VPPSPYPVAVTPG	APPLHQPSPAHAHPPPPPGTPPP	592
Db	534	VYRPPSPAPYPPQGPASHAPTTPYPAACP	PGPPPP--PCBSTQTRAPLTETPAFPP	590
Qy	593	AASLPQEPAPGAEGALYNASSAAHVNDTARA	ADLFVQMWGR	637
Db	591	AATGSQPEASNAEAGALYNASSAAHVVDTARA	ADLFVQMWGAR	635
RESULT 2				
A54862				
HSV-1 proteinase - human herpesvirus 1				
C;Species: human herpesvirus 1				
C;Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 09-Jul-2004				
C;Accession: A54862				
R;Darke, P.L.; Chen, E.; Hall, D.L.; Sardana, M.K.; Veloski, C.A.; Lafemina, R.				
J. Biol. Chem. 269, 18708-18711, 1994				
A;Title: Purification of active herpes simplex virus-1 protease expressed in E				
A;Reference number: A54862; MUID:94308117; PMID:8034621				
A;Accession: A54862				
A;Status: preliminary				
A;Molecule type: DNA				
A;Residues: 1-306 <DAR>				
A;Cross-references: UNIPROT:O69087; GB:L32018				
C;Superfamily: varicella-zoster virus gene 33 protein				
"Query Match 37.3%; Score 1262; DB 2; Length 306;				
Best Local Similarity 84.0%; Pred. No. 9.2e-59;				
Matches 252; Conservative 11; Mismatches 33; Indels 4; Gaps 1				
Qy	1	MASAEMLERLEAPLDRAVPVIVAGFLALYDSGDP	GELALDDPDIVRAALPENPLPINVD	60
Db	1	MAADAPGRNWSPLDRAVPVIVAGFLALYDSG	GELALDDPDIVRAALPDNDPINVD	60
Qy	61	HRARCEVGRVLAVVNDPRGPPFVVGLIACVQL	ERVLETAASAAIFERRGPALSREERLLYL	120
Db	61	HRAGCEVGRVLAVVNDPRGPPFVVGLIACVQL	ERVLETAASAAIFERRGPPLSREERLLYL	120
Qy	121	ITNYLPSVLSLTKRGDEVPDRTLFAHVALCAI	GRRLGTVITYDTSLDAAIAPFFHLDP	180
Db	121	ITNYLPSVLSATYKRLGGEAHPDRTLFAHVAL	CAIGRLGTVITYDGLDAAIAPFFHLSP	180
Qy	181	ATREGVRRAAEALALAGRTWAPGVEALHTLL	SLSTAVNNMLDRWSLVAERRRQAGTA	240
Db	181	ASREGARRLAAEALALSGRTWAPGVEALHTLL	SLSTAVNNMLDRWSLVAERRRQAGTA	240
Qy	241	GHTYLOASEKFKIWGASAPERGYKTTGAPG	AMDTSPAASVPAQ-----VAVRARSVAS	296
Db	241	GHTYLOASEKFKWGAEPVSPARGYKNGAP	ESTDIPPGSLAAAPQGDRCPIVQRQGVAS	300

Db 616 ASSQTVASASTGLDFGRDDADIFVQMMSAR 647

RESULT 5
S52216
viral proteinase - rabies virus
C;Species: rabies virus
C;date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 21-Aug-1998
C;Accession: S52216
R;Camacho, A.; Tabaro, E.
submitted to the EMBL Data Library, June 1994
A;Reference number: S52215
A;accession: S52216
A>Status: preliminary
A:Molecule type: genomic RNA
A;Residues: 1-522 <CAM>
A;Cross-references: EMBL:X79983
A;Note: the source is designated as pseudorabies virus
C;Superfamily: varicella-zoster virus gene 33 protein

Query Match 27.9%; Score 943; DB 2; Length 522;
Best Local Similarity 39.5%; Pred. No. 5.3e-42;
Matches 261; Conservative 61; Mismatches 155; Indels 184; Gaps 33;

QY 20 PIYVAGFLALYDSCDGEALDPDTVRALPPENPLPINVDHRCARCEVGRLAVVNDPRG 79
DB 3 PVTSGYLALYDR-DGEGELATREIVAAALPPAGPLFINIDHRACRDIGAVLVVDDRG 61
QY 80 PRFVGLIACVOLERVLETAASAAFERRRPALSRERLLYLITNYPVSLSLTKR-RGDG 138
DB 62 PFELGVNCPQLGAVLARAVAPDFGDMRP--SDERLLYLLSNYLPSSLSSRRLLPGD- 118
QY 139 VPEDRTLFARHALCAIGRLGTIVTDTSLDAIAIPRHLDPATREGVRREAEEALALA 198
DB 119 -ADETLFAHVLCVIGRRVGTIVVDSPERAVGFRELISAGSELLARAESP---A 174
QY 199 GRTWAEVEALHTLLSTAVNNMLDRNSLVAERRRQAGIAGHYLQAEEKPKWGAES 258
DB 175 ERVWHMSEDLTRALLSTAVNNMLLRDWELVAERREAGVRAHTYLQAT---MW---A 227
QY 259 APAPERGYKTGPAMDTSPTAASPVPQAVARQVASSSSSSFFPAPADMNPVASGAP 318
DB 228 GLLPKSG---ASPGA-----RAQ-----CGHGSP 248
QY 319 APPPGDGSLMTASHYNQLVTGSAPRHPPLTACGLPAAGTVAYCHPGAGPSPHYPPP 378
DB 249 AERTFGD--YIFVPAQYNQLVVNR-----PAP----- 275
QY 379 PAHPYPCMLFPAGSPLEAQIALVGAIDADRQAGGLPAAAGDHGIKGSAKRHRHEVQPE 438
DB 276 -----SLESOLGATVSA-AMDRHRRSPPSEP----RPPAKRRY----- 310
QY 439 YDCGRDEPDPRDPYYYPGEARPEPRPVDSRRARQASGPHTITALTVCATVTSLOBELAHMR 498
DB 311 |--DDYAADNAYYPGEAPP-----PRATSRAV---VSSLQRREISHLR 346
QY 499 A---RTHAP-YGYPYP-----VG-PYHHPHADTTETAQPPLY-PAKAVALYLPPI- 540
DB 347 AQHVRPYTTYAPAPQLQLPPGNAVGHPPHPH-HAAGALYPPMYAPQPGELHAPPSPVA 405
QY 541 HTAP--PGPPLSGAVPP-----PSYP-----PVAVTPGAPPL-----HQP-SPA 577
DB 406 HAVPALPGLPASRRCGFAHVPAQVVVFQQFVVVQAQFVAVPAAPPPRLRLQORHAPAAPV 465
QY 578 HAHP--PPPGPTPPPAASLPQEAPGACAGALVNASSAHNVNTDARAADLFVSQMGKS 636
DB 466 QAAAPRAPASAQPPVQASVSAPR-ETESPAPPIDASSAA---VACQGADIIVSQMQSQ 521
QY 637 R 637
DB 522 R 522

polyprotein - equine herpesvirus 2
N:Contains: capsid scaffold protein; proteinase/capsid protein
C:Species: equine herpesvirus 2
C:Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C:Accession: S55610; S55611
R:Telford, E.A.R.; Watson, M.S.; Aird, H.C.; Perry, J.; Davison, A.J.
J. Mol. Biol. 249, 520-528, 1995
A:Title: The DNA sequence of equine herpesvirus 2.
A:Reference number: S55594; MUID:95302501; PMID:7783207
A:Accession: S55610
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Molecule type: DNA
A:Residues: 1-643 <TEL>
A:Cross-references: UNIPROT:P52369; GB:U20824; NID:G695172; PIDN:AAC13804.1; PID:G695189
A:Experimental source: strain 86/67
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1995
C:Superfamily: varicella-zoster virus gene 33 protein
F:1-307/Product: proteinase/capsid protein #status predicted <MAT1>
F:308-643/Product: capsid scaffold protein #status predicted <MAT2>

Query Match 11.3%; Score 383.5; DB 2; Length 643;
Best Local Similarity 24.7%; Pred. No. 6.3e-13;
Matches 164; Conservative 86; Mismatches 248; Indels 167; Gaps 28;

Qy	13	PLPDRA-----VPIYVAGFLALYD-SGDPGELALDPTVRAALPENPLPINVDHRR 64
Db	26	PVPEAAASRASRAPVYLGGFVDFSPKDSRALYNPADVGAHLPLPGPIPLNVEHLQE 85
Qy	65	CEVGRVLAVNDPRGPFVGLIACVQLERVLE-TASAAIFERRGPAISREERLLYLITN 123
Db	86	AHVGTWTLGLHLTRYGLFCVAVITAEBEFTLLDRLCAASSVARTRADHHLPPNPTEMLHT 145
Qy	124	YLPVSLSLTKRRGDVEP---PDRTPFAHVALCAIGRLGTLITYDTSIDAAIAPFRHLD 179
Db	146	WLPELSLSS-IHPDALPGAKGGTPIFOHVALCMGQGRGTVAVGESLDWLISKTSLS 204
Qy	180	PATREGVRRBAEALAGRTWAPGEAL-----THTLLSTAVNNMLRDRSLVA 231
Db	205	PEERGAIAEGYA-----SPAPESLPPEHFTCSNEILMAKAIIDAGFIKNRLEILK 253
Qy	232	ERRRQAGIAGHTYLOASEKEFKINGAESAPAPERGYKTGAPGMDTSPASVPAQVAVRA 291
Db	254	TDRGVAEVKAFTYLKASVQ-----GLPANLDEVDSA-----RG 286
Qy	292	ROVASSSSSSSFPAPA---DMNPVSASGAPAPPPGPDGSGYLWIPASHY----- 336
Db	287	GEDPPHAIATTHPATDATMTMQQPFPAQA-PAGGCEDLISVPRSTFWMLQTNLDTM 345
Qy	337	NQLVTGQ-----SAPRHPLTACGLPAACTVAYGHGAGSPHYPPPPAH----- 381
Db	346	RTSLGQRFQPIDAPAAPAQLRVPPPAAPFV-HEGYYPAPVHFQVDVAAQYLPVPL 404
Qy	382	PYGMLPAGSPSLEAQIALVGAIAADRAGGLPAAAGD---HGIRGSAKRHRVEQPE 438
Db	405	PPFCAMPFPAPPLPDFY-----KYGGIPAPGYSVAHPAR-PGKRKE----- 445
Qy	439	YDCGRDEPRDF--PYYPGEARPEPRVDSSRAARQASGPHETITLVGAVTSLQQLAH 496
Db	446	-DC-----DDEFGPLPFGEI-----HKDVQSLSSIAALQSELKD 480
Qy	497	MRARTHAPYGPYPVGPYHHPHADTTPAQPRYPYA--KAVYLPFP-----HIAPGP-P 548
Db	481	IK---NSQPFQPLPQLOPQAPQAPQAPQAPQAPQAPQAPQAPQAPQAPQAPQAPQ 537
Qy	549	LSGAVPPSPYPVAVTGP-----APPLHQPSPAHAPPPPPPGTTPPAASLPQPEA 601

Db 538 ASGGVP-----CAPGPGVGCEPQAPQVTVTHNGHQAAPQAGGATGATAANVEQRQP 589
 Qy 602 PGAEA 606
 Db 590 EGGEA 594
 RESULT 9
 QQB3R
 BVRf2 (EC-RF3) protein - human herpesvirus 4 (strain B95-8)
 N:Contains: BVRf2 (EC-RF3a) protein
 C:Species: human herpesvirus 4, Epstein-Barr virus
 C>Date: 03-Aug-1984 #sequence revision 03-Aug-1984 #text_change 09-Jul-2004
 C/Accession: A03798; S33049; S33050; S33051
 C/Bankier, A.T.; Deininger, P.L.; Farrell, P.J.; Barrell, B.G.
 Mol. Biol. Med. 1, 21-45, 1983
 A:Title: Sequence analysis of the 17,166 bp EcoRI fragment C of B95-8 Epstein-Barr virus
 A/Reference number: A33065; MUID:85035713; PMID:6092825
 A/Accession: A03798
 A:Molecule type: DNA
 A/Residues: 1-605 <BAN>
 A/Cross-references: UNIPROT:P03234; EMBL:V01555; NID:G59074; PIDN:CAA24801.1; PID:gl3349
 R:Baer, R.; Bankier, A.T.; Biggin, M.D.; Deininger, P.L.; Farrell, P.J.; Gibson, T.J.; H
 Nature 310, 207-211, 1984
 A:Title: DNA sequence and expression of the B95-8 Epstein-Barr virus genome.
 A/Reference number: A03794; MUID:84270667; PMID:6087149
 A/Contents: annotation; protein coding region
 C:Superfamily: varicella-zoster virus gene 33 protein
 F:261-605/Product: BVRf2 (EC-RF3a) protein #status predicted <BVR>
 Query Match 11.0%; Score 372.5; DB 1; Length 605;
 Best Local Similarity 24.8%; Pred. No. 2.2e-12;
 Matches 166; Conservative 79; Mismatches 237; Indels 187; Gaps 28;
 Qy 17 RAVPTYVAGFLALYDSGDGPE--LALDPTVRAALPPENPLPINVDHRCCEVGRVLAVV 74
 Db 3 QAPSVYVCGFERPDA-PPKDACLDLDTVKSQLPKPLPLTVEHLPDAPVGSVFLY 61
 Qy 75 NDRPGPFVGLTIACVQLERVLETA-ASAAIFERRGPAISREERLLYITNLPVSLSSTK 133
 Db 62 QSRAGLFSAAITSGDFLSLLDSIYHDCIAQSORPLPREPKVEALHA-WLPSLSLAS- 119
 Qy 134 RRGDEVPPD-----RTLFAHVALCAIGRRLLTYTDTSLDAAIAFRRHLDPATR 183
 Db 120 ----LHPDIPTQTAGGKLSFDFHVSICALGRRRGTAVYGTDLAWLKHFSDLPSIA 174
 Qy 184 EGVREAAEALAGRTWAPGVEALTH-LLSTAVNNMLDRWSLVAERRRQAGIAGH 242
 Db 175 AQIENDANAAK-----RESCPEDHPLPLFKLIJAKADAGFLRNKRVETLRDQRGVANI PAE 230
 Qy 243 TYLOASEKFIWGABSAPEAPERYGTGAPGAMDTSPAASVPAPQAVARAVASSSSSSS 302
 Db 231 SYLKASD-----APDLQKPDKALQSPPASTD-----PATMLSGNAGEGATACGSA 277
 Qy 303 -----PPAPADMPVSAAGAPAPPDPPGSGYUWI 331
 Db 278 AAGQDLISVPRNTFTLLQTLNDLNKPPRTPLPYAAPLPFPFHOAIATAPSYGPGAGAVA 337
 Qy 332 PASHYNQLVTQSGAPRHPPLTACGLPAAGTVAGHPGA-----GPSPHYPPPAHP 382
 Db 338 PAGGY-----FTSPGGYYAGP-AGGDPGAFAMDHTVYHPHP-PPPAYFG 381
 Qy 383 YPGMLFAGPSPLEAQIAALVGAIAADROAGGLPAAAGDGHGIRSAKRRRHEVEQPEYDCG 442
 Db 382 LPG-LFGPPPPVPPYGGSHL-----RADYVPAPS-----RSNKRKRDPEDEEGG 426
 Qy 443 RDEPRDPFYPGEARPEPRPVDSSRAARQAAGFHTITLVAVTSLQOELAHMART- 501
 Db 427 -----LFFGE-----DATLYRKDIAGLSKSVNELQHTLQALRRRTL 462
 Qy 502 ---HAPYGPYPVGP-YHH-----PHADTETPAQPPRYPAKAYLPPPHIAPPGPPLS 550
 Db 463 SYGHTGVGVCQQGFCYTHSGFYGFQPHOSYEVR-----YYPHP----- 502


```

Db 432 -----ERDASSDQ-EDMS-FPGEA-----DHGKARKELKAHGRDNNNSGSDAKGD 476
Qy 479 TITAVGAVTSLOQE-LAHWEARTHAPVGPVPHPHADTETPAQPPRYPAKAVYL 537
Db 477 RYDDIRALQELKREMLAVRQIAPRALLAQAQLATPVASPTTTTSHQAEASEPQASTAAA 536
Qy 538 PPHIAPP-----GPPLSGAVPP 555
Db 537 ASPSTASSHGSKSAERGVWVNSCRVAPPLEAVNPP 571

RESULT 12
Q08B88
UL80 protein - human cytomegalovirus (strain AD169)
C:Contains: capsid assembly protein; viral proteinase (EC 3.4.21.-)
C:Species: human cytomegalovirus, human herpesvirus 5
A:Note: host Homo sapiens (man)
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
C:Accession: S09843; S51034; S51035
R:Chee, M.S.; Bankier, A.T.; Beck, S.; Bohnl, R.; Brown, C.M.; Cerny, R.; Horsnell, T.;
M.; Barrell, B.G.
Curr. Top. Microbiol. Immunol. 154, 125-169, 1990
A>Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus
A:Reference number: S09749; MUID:90269039; PMID:2161319
A:Accession: S09843
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-708 <CHE>
A:Cross-references: UNIPROT:P16753; EMBL:X17403; NID:959591; PIDN:CAA35353.1; PID:g17808
A:Note: possible protein-coding frames are given
R:Stevens, J.T.; Mapelli, C.; Tsao, J.; Hail, M.; O'Boyle II, D.; Weinheimer, S.P.; Dila
Eur. J. Biochem. 226, 361-367, 1994
A>Title: In vitro proteolytic activity and active-site identification of the human cytom
A:Reference number: S51034; MUID:9504793; PMID:8001553
A:Accession: S51034
A:Molecule type: protein
A:Residues: 110-131,'X',133-134 <ST2>
A:Note: assembly protein release site (Ala-256-Ser-257) and maturation site (Ala-643-Ser
C:Superfamily: cytomegalovirus capsid assembly proteinase
F:1-256/Product: viral proteinase #status predicted <PMAT>
F:336-708/Product: capsid assembly #status predicted <PMAT>
F:132/Active site: Ser #status experimental
F:256-257/Cleavage site: Ala-Ser (viral proteinase) #status experimental
F:643-644/Cleavage site: Ala-Ser (viral proteinase) #status experimental

Query Match 9.4%; Score 317.5; DB 1; Length 708;
Best Local Similarity 23.7%; Pred. No. 1.8e-09;
Matches 150; Conservative 67; Mismatches 196; Indels 219; Gaps 27;

Qy 20 PIYVAGFLALYD-SGDPGELALDPTVTRALPEN-----PLPINVDHRCARCEGRV 70
Db 13 PTVYVGFRLYDQSPDEAEELLPRDVVEHLWHAQGGQSPSLVALPLNINHDDTAVVGHV 72
Qy 71 LAVVNDPRGPFVGLIACVQLERVLETAASAIFERRGPALSS-REERLLVLTINYLPSVS 129
Db 73 AMQSVRDGLFCVGTSPFLFIVRRASEKSELVSGVPSPQDPQVFEFLSGSVAGLS 132
Qy 130 LSKRRGDEVP-----PDRTLFAHVALCAIGRRGLTIVTYDTSLDAAIAPFRHLDPAT 182
Db 133 LSS-RRCDVVEAATSLSGSETTFKHALCVSGRRRTGLAVYGRDPFWVTQRPDLTAAD 191
Qy 183 REGVREAAEALAGRTWAPVEAL-----THLLSTAVNNMRLDRWSLVAERR 235
Db 192 RDGLR-----AQWRCSGT---AVDASGDPFRSDSYGLLNSVDALYIRERLPLKYDQK 243
Qy 236 QAGIA-GHTYLOASEKFIKWAESAAPERGKVTGAPGMDTSPAASVPAPQVAVAROV 294
Db 244 LVGVTERESYKAS-----VSPEACDIK--AASAE 274
Qy 295 ASSSSSSSPAPADNMNPNVSASGAPPPPGDGSYLMIPASHYNQLVTGOSAPRHPRLTAC 354

```

```

Db 275 GDSRQAATPAAGARVPSSSPSPVEPP-----SPVQPP----- 308
Qy 355 GLPAAGCTVAVGHGAGSPSPHYPPPPA-----HPY-----PGMLFAGSPLEAQI 398
Db 309 ALPASPSVL---PAEGPPSPSPSEPAAASMSHPLSAAVPAATAPPGATVAGASPAVSSL 365
Qy 399 A-----ALVGAIAADROAGGLPAAAGDHGIRGSAKRRRHEVEQPEYDCGRDEPD 447
Db 366 AWPHDGVYLPKDAFFSILGASRSA--VPV----- 392
Qy 448 RDPYTPG-----EARPEPRPVDSRRRAQAQSPHETITALTALVCAVTSLOQELAHMART 501
Db 393 ---MYPGAVAAPSPASAPLPLPSYPAS-----YCAPVVGVDQLA---ARH 432
Qy 502 HAPYGVPPVGPVHHPHADTETPAQPPRYPAKAVYLPPIHAPGPPSLGAVPPSPVPV 561
Db 433 FADY-----VDPHY-----PGWGRRYE-----PAPSLHPSYP-----VPPP----- 463
Qy 562 AVTPGPAPPLHQPSAPAHAPPPPPPPGPTPPPA 593
Db 464 -----PSPAYYRRRDSFGGMDEPPS 483

RESULT 13
WZBEM4
gene 17 protein - saimiriine herpesvirus 1 (strain 11)
C:Species: saimiriine herpesvirus 1
A:Note: host Saimiri sciureus (common squirrel monkey)
C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 16-Jul-1999
C:Accession: H36807
R:Albrecht, J.
Submitted to the EMBL Data Library, January 1992
A:Description: Primary structure of the herpesvirus saimiri genome.
A:Reference number: A36806
A:Accession: H36807
A:Molecule type: DNA
A:Residues: 1-475 <ALB>
A:Cross-references: GB:X64346; NID:G60320; PIDN:CAA45641.1; PID:G60339
R:Albrecht, J.C.; Nicholag, J.; Biller, D.; Cameron, K.R.; Biesinger, B.; Newman, C.; Wi
J. Virol. 66, 5047-5058, 1992
A>Title: Primary structure of the herpesvirus saimiri genome.
A:Reference number: A37309; MUID:92333688; PMID:1321287
A:Contents: annotation; possible protein-coding frames
A:Note: neither amino acid nor nucleotide sequence is given
C:Genetics:
A:Gene: 17
C:Superfamily: varicella-zoster virus gene 33 protein

Query Match 8.9%; Score 301.5; DB 1; Length 475;
Best Local Similarity 21.8%; Pred. No. 8.1e-09;
Matches 135; Conservative 80; Mismatches 230; Indels 173; Gaps 20;

Qy 21 IYVAGF--LALYDSDGDPGELALDPTVTRALPENPLPINVDHRCARCEGRVLAHVNDPR 78
Db 4 VYVAGFVDVAVPKVDP-VLYNLDDVSKCLPTKPIPLNIEHLPESTIGHTGLYAVTH 62
Qy 79 GPFVGLIACVQLERVLETAASAIFERRGPALSSREERLLVLTINYLPSVLSLTKRGDE 138
Db 63 GVFCVGIHSEKFLHLTENLFSNSCAQATSKFLPYQPLLEMLHTLTPALSSSLCPTAQ 122
Qy 139 VPPDRTLFAHVALCAIGRRGLTIVTYDTSLDAAIAPFRHLDPATREGVREAAEALALA 198
Db 123 NAANTNMFORVSLCALGRRGTVAVYSMNLEDAISQFCSISOAEVENIVQDSKNVINSL 182
Qy 199 GRTWAPGVREALTHLLSTAVNNMRLDRWSLVAERRQAGIAGHTYLOASEKFIKWAESA 258
Db 183 PK---PVFNIDPHILMAKIDAGFIKDRQLQLKTDKGVAIKKLTYLKASE----- 230
Qy 259 APAPERGYKTGAPGMDTSPAASVPAPQVAVARQVAVSSSSSSSPFAPADNMNPNVSAGAP 318
Db 231 -----IGKPVTEIDIS-----EDMNQHI----- 248

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2005, 18:57:09 ; Search time 199 Seconds
(without alignments)
1841.777 Million cell updates/sec

Title: US-10-623-429-9

Perfect score: 3385

Sequence: 1 MASAEMRERLEAPLPDRAVP.....VNVDTARAADLFVSMQMSR 637

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt_02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3385	100.0	637	2 P89449	P89449 human herpe
2	3370.5	99.6	638	2 Q69527	Q69527 human herpe
3	3351.5	99.0	636	2 Q69104	Q69104 human herpe
4	2334	69.0	635	2 Q69087	Q69087 human herpe
5	2333	68.9	635	1 VP40_HHV11	P10210 human herpe
6	1820	53.8	329	2 P90341	P90341 human herpe
7	1729.5	51.1	608	2 Q806A3	Q806A3 cercopithec
8	1368.5	40.4	651	2 Q8JY97	Q8JY97 saimirine
9	1361	40.2	562	2 Q918P4	Q918P4 bovine herp
10	1131	33.4	247	2 O12271	O12271 human herpe
11	1079.5	31.9	646	1 VP40_EHV1B	P28936 equine herp
12	1079.5	31.9	646	2 Q6S6T7	Q6S6T7 equine herp
13	1079.5	31.9	646	2 AAS45919	Aas45919 equine he
14	1056	31.2	619	2 Q6X238	Q6X238 bovine herp
15	1056	31.2	619	2 AAR86135	Aar86135 bovine he
16	1056	31.2	647	2 Q39278	Q39278 equine herp
17	1010	29.8	621	1 VP40_BHV1C	P54817 bovine herp
18	1010	29.8	621	2 Q76PK0	Q76PK0 bovine herp
19	1010	29.8	621	2 CAB01600	Cab01600 bovine he
20	960	28.4	524	2 Q83417	Q83417 suid herpes
21	904	26.7	526	2 Q85027	Q85027 suid herpes
22	838.5	24.8	639	2 Q782R9	Q782R9 gallid herp
23	838.5	24.8	639	2 Q785G2	Q785G2 marek's dis
24	838.5	24.8	639	2 Q9YZ26	Q9YZ26 gallid herp
25	824.5	24.4	605	1 VP40_VZVD	P09286 varicella-z
26	824.5	24.4	605	2 Q6QCM2	Q6QCM2 human herpe
27	824.5	24.4	605	2 AAT07715	Aat07715 human her
28	824.5	24.4	605	2 AAT07791	Aat07791 human her
29	792	23.4	643	2 Q9DPR1	Q9dpr1 meleagrid h
30	787	23.2	663	2 Q9E6P2	Q9e6p2 meleagrid h
31	787	23.2	663	2 AAS01665	Aas01665 turkey he

RESULT 1

P89449	ID	P89449	PRELIMINARY;	PRT;	637 AA.
AC	P89449;				
DT	01-MAY-1997	(TReMBLrel. 03, Created)			
DT	01-MAY-1997	(TReMBLrel. 03, Last sequence update)			
DT	01-OCT-2003	(TReMBLrel. 25, Last annotation update)			
DE	Protease.				
GN	Name=UL26;				
OS	Human herpesvirus 2 (HHV-2) (Human herpes simplex virus 2).				
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;				
OC	Alphaherpesvirinae; Simplexvirus.				
OX	NCBI_TaxID=10310;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=HG52;				
RX	MEDLINE=87111457; PubMed=3027242;				
RA	McGeoch D.J., Moss H.W., McNab D., Frame M.C.;				
RT	"DNA sequence and genetic content of the HindIII 1 region in the short				
RT	unique component of the herpes simplex virus type 2 genome;				
RT	identification of the gene encoding glycoprotein G, and evolutionary				
RT	comparisons.;"				
RL	J. Gen. Virol. 68:19-38(1987).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=HG52;				
RX	MEDLINE=90278430; PubMed=2161906;				
RA	Everett R., Fenwick M.;				
RT	"Comparative DNA sequence analysis of the host shutoff genes of				
RT	different strains of herpes simplex virus: type 2 strain HG52 encodes				
RT	a truncated UL41 product.;"				
RL	J. Gen. Virol. 71:1387-1390(1990).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=HG52;				
RX	MEDLINE=92113549; PubMed=1662697;				
RA	McGeoch D.J., Cunningham C., McIntyre G., Dolan A.;				
RT	"Comparative sequence analysis of the long repeat regions and				
RT	adjoining parts of the long unique regions in the genomes of herpes				
RT	simplex viruses types 1 and 2.;"				
RL	J. Gen. Virol. 72:3057-3075(1991).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=HG52;				
RX	MEDLINE=92356101; PubMed=1322965;				
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;				
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane				
RT	protein with counterparts in other herpesviruses.;"				
RL	J. Gen. Virol. 73:2167-2171(1992).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=HG52;				
RX	MEDLINE=92356101; PubMed=1322965;				
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;				
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane				
RT	protein with counterparts in other herpesviruses.;"				
RL	J. Gen. Virol. 73:2167-2171(1992).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=HG52;				
RX	MEDLINE=92356101; PubMed=1322965;				
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;				
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane				
RT	protein with counterparts in other herpesviruses.;"				
RL	J. Gen. Virol. 73:2167-2171(1992).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=HG52;				
RX	MEDLINE=92356101; PubMed=1322965;				
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;				
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane				
RT	protein with counterparts in other herpesviruses.;"				
RL	J. Gen. Virol. 73:2167-2171(1992).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=HG52;				
RX	MEDLINE=92356101; PubMed=1322965;				
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;				
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane				
RT	protein with counterparts in other herpesviruses.;"				
RL	J. Gen. Virol. 73:2167-2171(1992).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=HG52;				
RX	MEDLINE=92356101; PubMed=1322965;				
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;				
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane				
RT	protein with counterparts in other herpesviruses.;"				
RL	J. Gen. Virol. 73:2167-2171(1992).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=HG52;				
RX	MEDLINE=92356101; PubMed=1322965;				
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;				
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane				
RT	protein with counterparts in other herpesviruses.;"				
RL	J. Gen. Virol. 73:2167-2171(1992).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=HG52;				
RX	MEDLINE=92356101; PubMed=1322965;				
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;				
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane				
RT	protein with counterparts in other herpesviruses.;"				
RL	J. Gen. Virol. 73:2167-2171(1992).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=HG52;				
RX	MEDLINE=92356101; PubMed=1322965;				
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;				
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane				
RT	protein with counterparts in other herpesviruses.;"				
RL	J. Gen. Virol. 73:2167-2171(1992).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=HG52;				
RX	MEDLINE=92356101; PubMed=1322965;				
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;				
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane				
RT	protein with counterparts in other herpesviruses.;"				
RL	J. Gen. Virol. 73:2167-2171(1992).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=HG52;				
RX	MEDLINE=92356101; PubMed=1322965;				
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;				
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane				
RT	protein with counterparts in other herpesviruses.;"				
RL	J. Gen. Virol. 73:2167-2171(1992).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=HG52;				
RX	MEDLINE=92356101; PubMed=1322965;				
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;				
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane				
RT	protein with counterparts in other herpesviruses.;"				
RL	J. Gen. Virol. 73:2167-2171(1992).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=HG52;				
RX	MEDLINE=92356101; PubMed=1322965;				
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;				
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane				
RT	protein with counterparts in other herpesviruses.;"				
RL	J. Gen. Virol. 73:2167-2171(1992).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=HG52;				
RX	MEDLINE=92356101; PubMed=1322965;				
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;				
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane				
RT	protein with counterparts in other herpesviruses.;"				
RL	J. Gen. Virol. 73:2167-2171(1992).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=HG52;				
RX	MEDLINE=92356101; PubMed=1322965;				
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;				
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane				
RT	protein with counterparts in other herpesviruses.;"				
RL	J. Gen. Virol. 73:2167-2171(1992).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=HG52;				
RX	MEDLINE=92356101; PubMed=1322965;				
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;				
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane				
RT	protein with counterparts in other herpesviruses.;"				
RL	J. Gen. Virol. 73:2167-2171(1992).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=HG52;				
RX	MEDLINE=92356101; PubMed=1322965;				
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;				
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane				
RT	protein with counterparts in other herpesviruses.;"				
RL	J. Gen. Virol. 73:2167-2171(1992).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=HG52;				
RX	MEDLINE=92356101; PubMed=1322965;				
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;				
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane				
RT	protein with counterparts in other herpesviruses.;"				
RL	J. Gen. Virol. 73:2167-2171(1992).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=HG52;				
RX	MEDLINE=92356101; PubMed=1322965;				
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;				
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane				
RT	protein with counterparts in other herpesviruses.;"				
RL	J. Gen. Virol. 73:2167-2171(1992).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=HG52;				
RX	MEDLINE=92356101; PubMed=1322965;				
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;				
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane				
RT	protein with counterparts in other herpesviruses.;"				
RL	J. Gen. Virol. 73:2167-2171(1992).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=HG52;				
RX	MEDLINE=92356101; PubMed=1322965;				
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;				
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane				
RT	protein with counterparts in other herpesviruses.;"				
RL	J. Gen. Virol. 73:2167-2171(1992).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=HG52;				
RX	MEDLINE=92356101; PubMed=1322965;				
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;				
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane				
RT	protein with counterparts in other herpesviruses.;"				
RL	J. Gen. Virol. 73:2167-2171(1992).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=HG52;				
RX	MEDLINE=92356101; PubMed=1322965;				
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;				
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane				
RT	protein with counterparts in other herpesviruses.;"				
RL	J. Gen. Virol. 73:2167-2171(1992).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=HG52;				
RX	MEDLINE=92356101; PubMed=1322965;				
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;				
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane				
RT	protein with counterparts in other herpesviruses.;"				
RL	J. Gen. Virol. 73:2167-2171(1992).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=HG52;				
RX	MEDLINE=92356101; PubMed=1322965;				
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;				
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane				
RT	protein with counterparts in other herpesviruses.;"				
RL	J. Gen. Virol. 73:2167-2171(1992).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=HG52;				
RX	MEDLINE=92356101; PubMed=1322965;				
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;				
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane				
RT	protein with counterparts in other herpesviruses.;"				
RL	J. Gen. Virol. 73:2167-2171(1992).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=HG52;				
RX	MEDLINE=92356101; PubMed=1322965;				
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;				
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane				
RT	protein with counterparts in other herpesviruses.;"				
RL	J. Gen. Virol. 73:2167-2171(1992).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=HG52;				
RX	MEDLINE=92356101; PubMed=1322965;				
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;				
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane				
RT	protein with counterparts in other herpesviruses.;"				
RL	J. Gen. Virol. 73:2167-2171(1992).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=HG52;				
RX	MEDLINE=92356101; PubMed=1322965;				
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;				
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane				
RT	protein with counterparts in other herpesviruses.;"				
RL	J. Gen. Virol. 73:2167-2171(1992).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=HG52;				
RX	MEDLINE=92356101; PubMed=1322965;				
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;				
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane				

```
DR EMBL; Z86099; CAB06750.1; -.
DR HSSP; Q69527; IAT3.
DR MEROPS; S21.001; -.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004252; F:serine-type endopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001847; Peptidase S21.
DR Pfam; PF00716; Peptidase S21; I.
DR PRINTS; PR00236; HSCAPSIDP40.
DR PROTEASE.
KW PROTEASE.
SQ
SEQUENCE 637 AA; 66940 MW; 9689D1F68C8DAB9C CRC64;
Query Match 100.0%; Score 3385; DB 2; Length 637;
Best Local Similarity 100.0%; Pred. No. 4.2e-145;
Matches 637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MASAEMRERLEAPLPDRAVPIYVAGFLALYDSGDPGELALDPDTVRAALPPENPLPINDV 60
Db 1 MASAEMRERLEAPLPDRAVPIYVAGFLALYDSGDPGELALDPDTVRAALPPENPLPINDV 60
QY 61 HRARCEVGRVLAVNDPRGPFVGLIACVQLSERVLETAASAAI FERRGPALSREERLLYL 120
Db 61 HRARCEVGRVLAVNDPRGPFVGLIACVQLSERVLETAASAAI FERRGPALSREERLLYL 120
QY 121 ITNYLPSVLSKTRRGDEVPPDRTLFAHVALCAIGRRIGTIIVTYDTSLDAAIAPFRHLDP 180
Db 121 ITNYLPSVLSKTRRGDEVPPDRTLFAHVALCAIGRRIGTIIVTYDTSLDAAIAPFRHLDP 180
QY 181 ATREGVRRERAAEALALAGRTWAPGVEALHTLLSTAVNNMLDRWSLVAERRRQAGIA 240
Db 181 ATREGVRRERAAEALALAGRTWAPGVEALHTLLSTAVNNMLDRWSLVAERRRQAGIA 240
QY 241 GHTYLOASEKFKIWGAEAPAPERGYKTGAPGAMDTSPAASVPAPQVAVRARQVASSSSS 300
Db 241 GHTYLOASEKFKIWGAEAPAPERGYKTGAPGAMDTSPAASVPAPQVAVRARQVASSSSS 300
QY 301 SSFPADNMNPFVSAGAPAPPPGDSYLMIPASHYNQLVTGOSAPRHPPLTACGLPAA 360
Db 301 SSFPADNMNPFVSAGAPAPPPGDSYLMIPASHYNQLVTGOSAPRHPPLTACGLPAA 360
QY 361 TVAYGHGAGSPHYPPPAHPYGMIFAGSPLEAQIAALVGAIAADRQAGGLPAAAGD 420
Db 361 TVAYGHGAGSPHYPPPAHPYGMIFAGSPLEAQIAALVGAIAADRQAGGLPAAAGD 420
QY 421 HGIRGSAKRHRHEVEQPEYDCGRDEPRDFYPYPGEARPEPRVDSRRARQAASGPHETI 480
Db 421 HGIRGSAKRHRHEVEQPEYDCGRDEPRDFYPYPGEARPEPRVDSRRARQAASGPHETI 480
QY 481 TALVGAVTSLQOELAHMRARTHAPYGPYPVGYHHPHADTETPAQPPRYPAKAVYLPPP 540
Db 481 TALVGAVTSLQOELAHMRARTHAPYGPYPVGYHHPHADTETPAQPPRYPAKAVYLPPP 540
QY 541 HTAPPGPPLSGAVPPPSYPVAVTPGAPPLHQPSPAHAHPPPPPGTPPPPAASLPQPE 600
Db 541 HTAPPGPPLSGAVPPPSYPVAVTPGAPPLHQPSPAHAHPPPPPGTPPPPAASLPQPE 600
QY 601 APGAEGALVNASSAAHVNVDARAADLFVSQMMGSR 637
Db 601 APGAEGALVNASSAAHVNVDARAADLFVSQMMGSR 637
RESULT 2
Q69527 ID Q69527 PRELIMINARY; PRT; 638 AA.
AC Q69527;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE UL26 protease.
OS Human herpesvirus 2 (HHV-2) (Human herpes simplex virus 2).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10310;
```

```

[1]
RN SEQUENCE FROM N.A.
RP O'Shannessy D.J., O'Donnell K.C., Hellmig B., Shabon U., O'Brien S.,
RA Chamberlain P., Holmes S., Abdel-Meguid S.S., Debouck C.M.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U49329; AAA92139.1; -.
DR PDB; IAT3; X-ray; A/B=1-247.
DR MEROPS; S21.001; -.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004252; F:serine-type endopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001847; Peptidase S21.
DR Pfam; PF00716; Peptidase S21; I.
DR PRINTS; PR00236; HSCAPSIDP40.
DR PROTEASE.
KW PROTEASE.
SQ
SEQUENCE 638 AA; 67028 MW; 992F230937BBAE89 CRC64;
Query Match 99.6%; Score 3370.5; DB 2; Length 638;
Best Local Similarity 99.7%; Pred. No. 1.9e-144;
Matches 636; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
QY 1 MASAEMRERLEAPLPDRAVPIYVAGFLALYDSGDPGELALDPDTVRAALPPENPLPINDV 60
Db 1 MASAEMRERLEAPLPDRAVPIYVAGFLALYDSGDPGELALDPDTVRAALPPENPLPINDV 60
QY 61 HRARCEVGRVLAVNDPRGPFVGLIACVQLSERVLETAASAAI FERRGPALSREERLLYL 120
Db 61 HRARCEVGRVLAVNDPRGPFVGLIACVQLSERVLETAASAAI FERRGPALSREERLLYL 120
QY 121 ITNYLPSVLSKTRRGDEVPPDRTLFAHVALCAIGRRIGTIIVTYDTSLDAAIAPFRHLDP 180
Db 121 ITNYLPSVLSKTRRGDEVPPDRTLFAHVALCAIGRRIGTIIVTYDTSLDAAIAPFRHLDP 180
QY 181 ATREGVRRERAAEALALAGRTWAPGVEALHTLLSTAVNNMLDRWSLVAERRRQAGIA 240
Db 181 ATREGVRRERAAEALALAGRTWAPGVEALHTLLSTAVNNMLDRWSLVAERRRQAGIA 240
QY 241 GHTYLOASEKFKIWGAEAPAPERGYKTGAPGAMDTSPAASVPAPQVAVRARQVASSSSS 299
Db 241 GHTYLOASEKFKIWGAEAPAPERGYKTGAPGAMDTSPAASVPAPQVAVRARQVASSSSS 300
QY 300 SSFPADNMNPFVSAGAPAPPPGDSYLMIPASHYNQLVTGOSAPRHPPLTACGLPAA 359
Db 300 SSFPADNMNPFVSAGAPAPPPGDSYLMIPASHYNQLVTGOSAPRHPPLTACGLPAA 360
QY 360 GTVAYGHGAGSPHYPPPAHPYGMIFAGSPLEAQIAALVGAIAADRQAGGLPAAAG 419
Db 360 GTVAYGHGAGSPHYPPPAHPYGMIFAGSPLEAQIAALVGAIAADRQAGGLPAAAG 420
QY 420 DHGIRGSAKRHRHEVEQPEYDCGRDEPRDFYPYPGEARPEPRVDSRRARQAASGPHET 479
Db 420 DHGIRGSAKRHRHEVEQPEYDCGRDEPRDFYPYPGEARPEPRVDSRRARQAASGPHET 480
QY 480 ITALVGAVTSLQOELAHMRARTHAPYGPYPVGYHHPHADTETPAQPPRYPAKAVYLPPP 539
Db 480 ITALVGAVTSLQOELAHMRARTHAPYGPYPVGYHHPHADTETPAQPPRYPAKAVYLPPP 540
QY 540 PHIAPPGPPLSGAVPPPSYPVAVTPGAPPLHQPSPAHAHPPPPPGTPPPPAASLPQPE 599
Db 540 PHIAPPGPPLSGAVPPPSYPVAVTPGAPPLHQPSPAHAHPPPPPGTPPPPAASLPQPE 600
QY 600 EAPGAEGALVNASSAAHVNVDARAADLFVSQMMGSR 637
Db 600 EAPGAEGALVNASSAAHVNVDARAADLFVSQMMGSR 638
RESULT 3
Q69104 ID Q69104 PRELIMINARY; PRT; 636 AA.
AC Q69104;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
```

DE Protease (Fragment).
GN Name=ICP35;
OS Human herpesvirus 2 (HHV-2) (Human herpes simplex virus 2).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCB1_TaxID=10310;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=G;
RA Steffy K.K.S., Schoen S., Chen C.-M.;
RT "Nucleotide sequence of the herpes simplex virus type 2 gene encoding
RT the protease and capsid protein ICP35.";
RL J. Gen. Virol. 0:0-0(1995).
DR EMBL; L37443; AAA67703.1; -.
DR MEROPS; S21.001; -.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004252; F:serine-type endopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001847; Peptidase_S21.
DR Pfam; PF00716; Peptidase_S21; 1.
DR PRINTS; PR00236; HSCVAPSIDP40.
KW Protease.
FT NON_TER
FT NON_TER
SQ SEQUENCE 636 AA; 66785 MW; 817FF5814F7A48C3 CRC64;
Query Match 99.0%; Score 3351.5; DB 2; Length 636;
Best Local Similarity 99.4%; Pred. No. 1.3e-143;
Matches 633; Conservative 1; Mismatches 2; Indels 1; Gaps 1;
QY 1 MASAEMRELEAPLPDRAVPIYVAGFLALYDSGDELALDPDTVRAALPPENPLPINVD 60
DB 1 MASAEMRELEAPLPDRAVPIYVAGFLALYDSGDELALDPDTVRAALPPENPLPINVD 60
QY 61 HRACEVGRVLAVVNDPRGPFVGLIACVOLERVLETAASAAIFERRGPALSREERLLYL 120
DB 61 HRACEVGRVLAVVNDPRGPFVGLIACVOLERVLETAASAAIFERRGPALSREERLLYL 120
QY 121 ITNYPVSLSLTKRRGDEVPPDRTLFAHVALCAIGRLGTIVYDTSLDAAIAPFRHLDP 180
DB 121 ITNYPVSLSLTKRRGDEVPPDRTLFAHVALCAIGRLGTIVYDTSLDAAIAPFRHLDP 180
QY 181 ATREGVRRREAAEALAGRTWAPGVEALHTLLSTAVNNMLDRWSLVAERRRQAGIA 240
DB 181 ATREGVRRREAAEALAGRTWAPGVEALHTLLSTAVNNMLDRWSLVAERRRQAGIA 240
QY 241 GHTYLQASEKFKIWGAEAPAPERGKYGAPGAMDTSPAASVPAQVAVRQVASSSS 300
DB 241 GHTYLQASEKFKIWGAEAPAPERGKYGAPGAMDTSPAASVPAQVAVRQVASSSS 300
QY 301 SSFPAPADMPVSGAPAPPPGDSGLYLPASHYNQLVTGQSAAPHPLTACGLPAAG 360
DB 301 SSFPAPADMPVSGAPAPPPGDSGLYLPASHYNQLVTGQSAAPHPLTACGLPAAG 360
QY 361 TVAYGHGAGSPSPHYPPPAHPYPGMLFAGSPLEAIAALVGAIAADROAGGLPAAAGD 420
DB 361 TVAYGHGAGSPSPHYPPPAHPYPGMLFAGSPLEAIAALVGAIAADROAGGLPAAAGD 420
QY 421 HGIRGSAKRHRHEVEQPEYDCGRDEPDRTFPYVGEARPEPRVDSRAARQASGPHETI 480
DB 421 HGIRGSAKRHRHEVEQPEYDCGRDEPDRTFPYVGEARPEPRVDSRAARQASGPHETI 480
QY 481 TALVGVATSLQQLAHMRARHAPYGYPPVGVPHPHADTETPAOPPRYPAKAVLPPP 540
DB 481 TALVGVATSLQQLAHMRARHAPYGYPPVGVPHPHADTETPAOPPRYPAKAVLPPP 540
QY 541 HIAPPGPLSGAVPPSPYVAVTPGAPPLHQPSPAHAPHPPPPPGTPPPPAASLPQPE 600
DB 541 HIAPPGPLSGAVPPSPYVAVTPGAPPLHQPSPAHAPHPPPPPGTPPPPAASLPQPE 600
QY 601 APGAEGALVNASSAAHVNDTARAADLFVQMMGSR 637
DB 601 APGAEGALVNASSAAHVNDTARAADLFVQMMGSR 637

RESULT 4
Q69087 PRELIMINARY; PRT; 635 AA.
AC Q69087;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE UL26.
OS Human herpesvirus 1 (HHV-1) (Human herpes simplex virus 1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCB1_TaxID=10298;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94308117; PubMed=8034621;
RA Darke P.L., Chen E., Hall D.L., Sardana M.K., Veloski C.A.,
RA Lafemina R.L., Shafer J.A., Kuo L.C.;
RT "Purification of active herpes simplex virus-1 protease expressed in
RT Escherichia coli.";
RL J. Biol. Chem. 269:18708-18711(1994).
DR EMBL; L32018; AAA5828.1; -.
DR PIR; A54862; A54862.
DR HSSP; Q69527; 1AT3.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004252; F:serine-type endopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001847; Peptidase_S21.
DR Pfam; PF00716; Peptidase_S21; 1.
DR PRINTS; PR00236; HSCVAPSIDP40.
KW Protease.
FT CHAIN 1 247 HSV-1 protease.
FT CHAIN 253 610 ICP35 assembly protein.
SQ SEQUENCE 635 AA; 66454 MW; 3C7768EE1496FDBF CRC64;
Query Match 69.0%; Score 2334; DB 2; Length 635;
Best Local Similarity 74.0%; Pred. No. 9.9e-98;
Matches 477; Conservative 30; Mismatches 120; Indels 18; Gaps 7;
QY 1 MASAEMRELEAPLPDRAVPIYVAGFLALYDSGDELALDPDTVRAALPPENPLPINVD 60
DB 1 MASAEMRELEAPLPDRAVPIYVAGFLALYDSGDELALDPDTVRAALPPENPLPINVD 60
QY 61 HRACEVGRVLAVVNDPRGPFVGLIACVOLERVLETAASAAIFERRGPALSREERLLYL 120
DB 61 HRACEVGRVLAVVNDPRGPFVGLIACVOLERVLETAASAAIFERRGPALSREERLLYL 120
QY 121 ITNYPVSLSLTKRRGDEVPPDRTLFAHVALCAIGRLGTIVYDTSLDAAIAPFRHLDP 180
DB 121 ITNYPVSLSLTKRRGDEVPPDRTLFAHVALCAIGRLGTIVYDTSLDAAIAPFRHLDP 180
QY 181 ATREGVRRREAAEALAGRTWAPGVEALHTLLSTAVNNMLDRWSLVAERRRQAGIA 240
DB 181 ASREGARRLAAEALAGRTWAPGVEALHTLLSTAVNNMLDRWSLVAERRRQAGIA 240
QY 241 GHTYLQASEKFKIWGAEAPAPERGKYGAPGAMDTSPAASVPAQVAVRQVASSSS 296
DB 241 GHTYLQASEKFKIWGAEAPAPERGKYGAPGAMDTSPAASVPAQVAVRQVASSSS 296
QY 297 SSSSSFPAPADMPVSGAPAPPPGDSGLYLPASHYNQLVTGQSAAPHPLTACGL 356
DB 297 SSSSSFPAPADMPVSGAPAPPPGDSGLYLPASHYNQLVTGQSAAPHPLTACGL 356
QY 301 P-----PVLPMNVPVTSPTAPAPPGDSGLYLPASHYNQLVAGHAAPQPHSAFGF 354
DB 301 P-----PVLPMNVPVTSPTAPAPPGDSGLYLPASHYNQLVAGHAAPQPHSAFGF 354
QY 357 P-AAGTVAGHPGAGSPHYPPPAHPYPGMLFAGSPLEAIAALVGAIAADROAGGLP 415
DB 355 PAAAGSVAYGHPGAGLSQHPYVHVHAYQYGVLFSGSPLEAIAALVGAIAADROAGGP 414
QY 416 AAGDGHGIRGSAKRHRHEVEQPEYDCGRDEPDRTFPYVGEARPEPRVDSRAARQASG 475
DB 415 -AAGDPGVNGSGKRRRYEAGPSSEYCDQDEPDADYPPYVGEARGA PRGVDSRAARHSPG 473
QY 476 PHETITLVGAVTSLQQLAHMRARHAPYGYPPVGVPHPHADTE-TPAQPPRYPAKA 534

RC STRAIN=HG52;
RX MEDLINE=871111457; PubMed=3027242;
RA McGeoch D.J., Moss H.W., McNab D., Frame M.C.;
RT "DNA sequence and genetic content of the HindIII 1 region in the short
RT unique component of the herpes simplex virus type 2 genome;
RT identification of the gene encoding glycoprotein G, and evolutionary
RT comparisons."; J. Gen. Virol. 68:19-38(1987).
RL [2]
RN SEQUENCE FROM N.A.
RC STRAIN=HG52;
RX MEDLINE=90278430; PubMed=2161906;
RA Everett R., Fenwick M.;
RT "Comparative DNA sequence analysis of the host shutoff genes of
RT different strains of herpes simplex virus: type 2 strain HG52 encodes
RT a truncated UL41 product."; J. Gen. Virol. 71:1387-1390(1990).
RL [3]
RN SEQUENCE FROM N.A.
RC STRAIN=HG52;
RX MEDLINE=92113549; PubMed=1662697;
RA McGeoch D.J., Cunningham C., McIntyre G., Dolan A.;
RT "Comparative sequence analysis of the long repeat regions and
RT adjoining parts of the long unique regions in the genomes of herpes
RT simplex viruses types 1 and 2."; J. Gen. Virol. 72:3057-3075(1991).
RL [4]
RN SEQUENCE FROM N.A.
RC STRAIN=HG52;
RX MEDLINE=92356101; PubMed=1322965;
RA Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;
RT "A novel herpes simplex virus gene (UL49A) encodes a putative membrane
RT protein with counterparts in other herpesviruses."; J. Gen. Virol. 73:2167-2171(1992).
RL [5]
RN SEQUENCE FROM N.A.
RC STRAIN=HG52;
RX Dolan A.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RA EMBL; Z86099; CAB06751.1; -;
RQ SEQUENCE 329 AA; 33836 MW; 59B0B428E6345186 CRC64;
DR

Query Match	53.8%;	Score 1820;	DB 2;	Length 329;	
Best Local Similarity	100.0%;	Pred. No. 8e-75;			
Matches 329;	Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	309	MNPVSASGAPAPPPGGDGYLWTPASHYNOLVTGOSAPRHPPPTACGLPAAGTVAYGHG	368		
Db	1	MNPVSASGAPAPPPGGDGYLWTPASHYNOLVTGOSAPRHPPPTACGLPAAGTVAYGHG	60		
Qy	369	AGPSPHYPPPPAHFYPCGMLFAGSPLEAQTAALVGAIAADRQAGGLPAAAGDHGIRSAK	428		
Db	61	AGPSPHYPPPPAHFYPCGMLFAGSPLEAQTAALVGAIAADRQAGGLPAAAGDHGIRSAK	120		
Qy	429	RRRHEVEQPYDCGRDEPDRDPFYYPGEARPEPRPVDSSRAAQSGPHTTITALVGAVT	488		
Db	121	RRRHEVEQPYDCGRDEPDRDPFYYPGEARPEPRPVDSSRAAQSGPHTTITALVGAVT	180		
Qy	489	SLOQELAHMRATHAPYGPYPVPGYPHYHHPHADTETPAQPPRYPAKAVYLPHPPIAPGPP	548		
Db	181	SLOQELAHMRATHAPYGPYPVPGYPHYHHPHADTETPAQPPRYPAKAVYLPHPPIAPGPP	240		
Qy	549	LSCGAVPPSPYPVAVTGPAPPLHQPSPAHAPPPPPPGTTPPPAAQLPQEPAPGAAGA	608		
Db	241	LSCGAVPPSPYPVAVTGPAPPLHQPSPAHAPPPPPPGTTPPPAAQLPQEPAPGAAGA	300		
Qy	609	LVNASSAAHVNVDARAADLVFSQMMGSR	637		
Db	301	LVNASSAAHVNVDARAADLVFSQMMGSR	329		

RESULT 7
Q806A3

Q806A3	PRELIMINARY;	PRT;	608 AA.
Q806A3;			
AC	01-JUN-2003 (TrEMBLrel. 24, Created)		
DT	01-JUN-2003 (TrEMBLrel. 24, Last sequence update)		
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)		
DE	Capsid maturation protease.		
GN	Name=UL26;		
OS	Cercopithecine herpesvirus 1 (CeHV-1) (Simian herpes B virus).		
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;		
OC	Alphaherpesvirinae; Simplexvirus.		
OX	NCBI_TaxID=10325;		
OX	[1]		
RP	SEQUENCE FROM N.A.		
RP	MEDLINE=22607624; PubMed=12721804;		
FX	Ohsewa K., Black D.H., Sato H., Rogers K., Eberle R.;		
RA	"Sequence and genetic arrangement of the UL region of the monkey B		
RT	virus (Cercopithecine herpesvirus 1) genome and comparison with the UL		
RT	region of other primate herpesviruses.";		
RT	Arch. Virol. 148:989-997(2003).		
RL	[2]		
RL	SEQUENCE FROM N.A.		
RP	STRAIN=E2490.		
RC	MEDLINE=22628476; PubMed=12743273;		
FX	Pereyagina L., Zhu L., Zurbukhien H., Mills R., Borodovsky M.,		
RA	Hilliard J.K.;		
RA	"Complete sequence and comparative analysis of the genome of herpes B		
RT	virus (Cercopithecine herpesvirus 1) from a rhesus monkey.";		
RT	J. Virol. 77:6167-6177(2003).		
RL	EMBL; AB096160; BAC58065.1;		
DR	EMBL; AF533768; AAP41443.1;		
DR	HSSP; P09286; IVZV.		
DR	GO; GO:0008233; F:peptidase activity; IEA.		
DR	GO; GO:0004252; F:serine-type endopeptidase activity; IEA.		
DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.		
DR	InterPro; IPR001847; Peptidase S21.		
DR	Prfam; PF00716; Peptidase S21; I.		
DR	PRINTS; PR00236; HSCVAPSIDP40.		
KW	Protease.		
QY	SEQUENCE 608 AA; 62683 MW; 16BA7A1C54FCAA3C CRC64;		
Query Match	51.1%; Score 1729.5; DB 2; Length 608;		
Best Local Similarity	59.4%; Pred. No. 1.7e-70;		
Matches 401; Conservative 28; Mismatches 143; Indels 101; Gaps 24;			
QY	1 MASAEMRLEAPLPDRAVPPIYVAGFLALYDSDGDELALDPTVRAALPPENPLPINVD	60	
Db	1 MGPAADPEGPGDADRPPIYVAGVLYXGSDGSELALDPTVRAALPPAGLAINVD	60	
QY	61 HRAECVGRVLAIVNDPRGPFVFFGLIACVQLERVLETAASAALFERRGPALSREERLLYL	120	
Db	61 HRAECVGRVLAIVNDPRGPFVFFGLIACVQLERVLETAASAALFERRGPALSREERLLYL	120	
QY	121 ITNYLPSVLSLTVRGDEVPDRTLFAHVALCAIGRLGTIVTYDTSLSDAIAFPFRLDP	180	
Db	121 VTNLYPSVLSLTVRLEEGETPDNTFLVHVALCAIGRLGTIVTYDTSLSDAIAFPFRLAP	180	
QY	181 ATREGVRRAAEALALAGRTWAPGVEALHTLLSTAVNNMLDRWSLVAERRROAGIA	240	
Db	181 DSREILURRANAELGLGGRVTPGAEALTRALLSTAVNNMLDRWSLVAERRRQNGIA	240	
QY	241 GHTYLQASEKFTIWAESAAPAPERG-YKTGAPAMDTSPAASVPAQVAVRARQVASSSS	299	
Db	241 GHTYLQASEKFLWAGPGDSPDGVYKRRSSCATDPGHIDASAPR-----AHPD	291	
QY	300 SSSFPADAMNVV-SASGAPAPPPDGGSYLWTPASHYQLVLTGQSAPRHPPLTACGLPA	358	
Db	292 PSDVP-EAEMNPAAPPAGAPTAKPPDGGSYLWTPAAHYNLQVLVAG-----HPPAPVFGAPA	345	
QY	359 AGTVAYGHPGAGSPHYPPPAHPYPGMLFAGSPLEAQIALVGLAIAADROAGGAPAA	418	
Db	346 P-PPAEGAYAPHHH-----AGVYGGVFFGPGSPLEAQIALVGLAIAADROGPGSP---	396	
QY	419 GDHGIRGSKRRHHVEQPEYDCGRDEPDRTFFYPGEARPEPRPV-DS-----RRAARQA	473	

```

Db 397 -----SGKRRYAEER---DCYDDRD-DAPYFGEARAPPVRVPUSGGRGRAPPSA 445
QY 474 -SGPHETITALVAVTSLQOELAHMARHATHPYG-PY-----PPVGPY----- 514
Db 446 GTGASETIALVAVTSLQOELAHMARARA-APYQPYAQAARPALGVAGEPPAPQVAQQWE 504
QY 515 -----HHPHADTETPAQPPRYPAKAVYLPPIHAIAPGPPPLSGAVPPSPYVAV-TP 565
Db 505 APAAGALAHHP-----PPPRH-----QPHATPANVEL-----VFGPQCAVALAAQ 545
QY 566 GPAPPLHOPSPAHAPPPPPGPTPPPAASLPQPEAPGAEA-GALVNASSAAHVNVDTAR 624
Db 546 GPAPGAAPGGQCA-----PAAGAAPPAA-----EAGETEAVGALVNASSAATHVDVDVGR 595
QY 625 AADLFVSQMMGR 637
Db 596 AADLFVSQMMGR 608

RESULT 8
Q8JY97 PRELIMINARY; PRT; 651 AA.
AC Q8JY97;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Protease.
OS Saimirine herpesvirus 1 (strain MV-5-PSL) (SaHV-1) (Marmoset
herpesvirus).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae.
OX NCBI_TaxID=10353;
RN [1]
RP SEQUENCE FROM N.A.
RA Brehm M.A., Black D.H., Ritchey J.W., Eberle R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY095366; AAM22798.1; -.
DR HSSP; Q69527; IAT3.
DR MEROPS; S21.001; -.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004252; F:serine-type endopeptidase activity; IEA.
DR GO; GO:0004508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001847; Peptidase_S21.
DR Pfam; PF00716; Peptidase S21; 1.
DR PRINTS; PRO0236; HSCVAPSIDP40.
KW Protease.
SQ SEQUENCE 651 AA; 68104 MW; CF62E03E2EFF3BB9 CRC64;

Query Match 40.4%; Score 1368.5; DB 2; Length 651;
Best Local Similarity 49.1%; Pred. No. 3.4e-54;
Matches 340; Conservative 54; Mismatches 182; Indels 117; Gaps 25;

QY 8 ERLEAPLPRAVP--LYVAGFLALYDSGDPGELALDPDTVRAALPENPLPINVDHARC 65
Db 11 EREDGAVDRTPTRMVYAGFLALYDSDDAEVLVADPVVAASLPAPPLPINDKTRC 70
QY 66 EVGRVLAVNDPRGPFVGLIACVQLERVLETAASAAIFERRGPALSREERLLYLITNYL 125
Db 71 SVQVLAVNDARGPFFLGVAGQLERHLEAASGDI FGRGPPLTRERLLYLITNYL 130
QY 126 PSVSLTKRGDEVP---PDRTLFAHVALCATGRRLGTIVTDTSLDAAIAFRHLDPAT 182
Db 131 PSVSLSTRDDSGAPVGN--LFAHVALCATGRRLGTIVTDTVMTEGAVAPFRRLGLAS 188
QY 183 RGVREAAEAELAGRTWAPVEALTHLLSTAVNNMLRDRWSLVAERRRQAGIAGH 242
Db 189 RGVLDAVAEAVSGERTWAPGARALTTLTATAVNNMLRDRWSVVAERRRQAGIAGH 248
QY 243 TYLQASEKFKINGASAPAPRGYKTCGAPCAMDTS--PAASVPAPQVAVRQVASS--- 297
Db 249 TYLQASAKFGL--SDASP-----RADAESAADQRHIRADGTQAPESAARERRHPSLR 300

```

```

QY 298 -----SSSSFPAPADMNVFSASGAPA-PPPPGDSYLIWIPASHYNQL 339
Db 301 PPSPPCDPSRI DCGSLVFSAGPMVNP-SAAGSPCLPKPPGDSYLIWIPASHYNHL 359
QY 340 VTGQ--SAPRHPLTACGLPAAGTVAYGHGPGAGSPHYPPPPAHYPGMLFAGSPLEAQ 397
Db 360 VSSQAATAPHLPIYSAHQVAG--YGPWGVVPA--YVPVPA-----AGSPLEAQ 407
QY 398 IALVGAIAADROAGGL-----PAAAGDHGIRGSAKRHRHEVEQPEYD-----CGRDEPDR 449
Db 408 IALVGAIAADROAGGL-----PAAAGDHGIRGSAKRHRHEVEQPEYD-----CGRDEPDR 464
QY 450 FPTYEGE-----ARPEPRVDSR--RAARQASG-PH--ETITLVGAVTSLQOELAHMAR 500
Db 465 -PYTPGEVSGGGEGPGRDRRFGTATRSAAATPHSNETIALVGAVALQOELTHLSY 523
QY 501 THAPYGPYPVPGYHHPHADTETPAQPPRYPAKAVYLPPIHAIAPGPPPLSGAVPPSPY 560
Db 524 QHGAF-----VPOSAAGVWTPRPYFAP-----AAAAQASHQL 556
QY 561 VAVTGPAPPLHOPSPAHAPPPPPGPTPPPAASL-----PQPEAPGA-----EAG 607
Db 557 QQAQFSCAPVTQTTPPAQVVPAAALPAPPAAPVQSLGVGVGAPMEFRAGDAADVASLEAD 616
QY 608 A---LVNASSAAHVNVDTARADLFVSQMMGR 637
Db 617 APPLINASCTRTVTDANRASDAFVAQMMGDR 649

RESULT 9
Q918P4 PRELIMINARY; PRT; 562 AA.
AC Q918P4;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE UL26.
OS Bovine herpesvirus 2.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10295;
RN [1]
RP SEQUENCE FROM N.A.
RA May J.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF410478; AAL05039.1; -.
DR HSSP; Q69527; IAT3.
DR MEROPS; S21.001; -.
DR GO; GO:0004252; F:serine-type endopeptidase activity; IEA.
DR GO; GO:0004508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001847; Peptidase_S21.
DR Pfam; PF00716; Peptidase S21; 1.
DR PRINTS; PRO0236; HSCVAPSIDP40.
SQ SEQUENCE 562 AA; 59809 MW; 1E783D9743D90CBD CRC64;

Query Match 40.2%; Score 1361; DB 2; Length 562;
Best Local Similarity 51.1%; Pred. No. 6.5e-54;
Matches 327; Conservative 47; Mismatches 154; Indels 112; Gaps 21;

QY 17 RAVPIYVAGFLALYDSGDPGELALDPDTVRAALPENPLPINVDHRCCEVGRVLAVND 76
Db 16 RAVPIYVAGFLALYEGGDELALPREVVSSALPSGVPVINVDRHRAQCEVGRVLTVDD 75
QY 77 PRGPFVGLIACVQLERVLETAASAAIFERRGPALSREERLLYLITNYLPSVSLTKR-- 134
Db 76 PRGPFVGLVSCAQLEEAASALFERRGPPLSREERLLYLITNYLPSVLSRRLE 135
QY 135 RGVDEVPDRTLFAHVALCATGRRLGTIVTDTSLDAAIAPEHLDPATREGVREAAAE 194
Db 136 PGDTSWED--LFRHVALCVIGRLGTIVTDTSVISGAVAPFOHLSRDAAIRAEAAAS 193
QY 195 LALAGRTWAPVEALTHLLSTAVNNMLRDRWSLVAERRRQAGIAGTLYQASEKFKTW 254

```


Db	194	LALRDREWAPCGDALARATLLSTAVNVSMMLRDWSLVSRRRQAGIAGHTYTLQASFAPIW	255
Qy	255	GAESAPAPERGYKTCAGAMDTPSAASVPAPQVAVRARQVASSSSSSSF-----PAP-AD	308
Db	254	GAGPTCAPSGGCKEHA-----RSPRATA-----SPEAGSGTFSVQVIGAPLPTAH	300
Qy	309	MNIVSASGAPAPP---PPGDSYLIWIPASHYNQLVTGQSAHPHPPLTACGLPAAGTVAYG	365
Db	301	MN-----CGPPAAPQARPSGDDSYLIWIPAAHYNQLVAAQ---PNHCQOOHMQFP---RVSHG	351
Qy	366	HPGAGPSPHYPPPAHPYP-----GWLPAAGSPLEAQIAALVCAIADRQAGGLPAAAGDH	421
Db	352	GPMGPPGVGH-----PIYTPFVGLGAVAEVSPLETOIAALVCAIADRQATDRNAA---EL	405
Qy	422	GIRGSAKRHRHEVQPEYDCGRDEP-----DRDFPYPYGEARPEPRPVDSSRAARQASGPH	477
Db	406	RSRGGGKRRN-----DYDDDDGSPPRYVGHGRDVPYYPGEGAPIRRAPEQRPA---VSPD	458
Qy	478	ETITALVGAVTSLQOELAHMARTHAPYGPYPVGYHHPHADTTPAQPPRYPAKAVTL	537
Db	459	DTITALTGAVSSLOQELAHMRSQVAVCAVP-----TTAPAPQP-----	496
Qy	538	PPPHIAPGPPSLGAVPPSPYPVAVTPGPAPPLHQPSPAHARHPPPPGTTPPAAASLP	597
Db	497	-----LPPSSQSVQATTQG-----HQP-----QIPQTTSATPQPVAABP	531
Qy	598	QPEAPGAEGALVNASSAAHVNVDTRAADLFVSQMMGR	637
Db	532	PPQ-----VDASGVAKVDVDARRAADLFVAHMMGR	562

RESULT 10

ID	O12271	PRELIMINARY;	PRT;	247	AA.
AC	O12271				
DT	01-JUN-1997	(TrEMBLrel. 04, Created)			
DT	01-JUL-1997	(TrEMBLrel. 04, Last sequence update)			
DT	01-OCT-2003	(TrEMBLrel. 25, Last annotation update)			
DE		Viral core protein.			
OS		Human herpesvirus 1 (HHV-1) (Human herpes simplex virus 1).			
OC		Viruses; dsDNA viruses, no RNA stage; Herpesviridae;			
OC		Alphaherpesvirinae; Simplexvirus.			
NCBI_TaxID		10298;			
RN		[1]			
RP		SEQUENCE FROM N.A.			
RX		MEDLINE=97433272; PubMed=9288912;			
RA		Apeler H., Gottschalk U., Guntermann D., Hansen J., Maessen J.,			
RA		Schmidt E., Schneider K., Schneider M., Ruebsamen-Waigmann H.,			
RT		"Expression of natural and synthetic genes encoding herpes simplex			
RT		virus 1 protease in Escherichia coli and purification of the			
RT		protein.";			
RL		Eur. J. Biochem. 247:890-895 (1997).			
DR	EMBL;	Z97070; CAB09805.1; --			
DR	HSSP;	Q69527; IAT3.			
DR	GO;	GO:0008233; F:Peptidase activity; IEA.			
DR	GO;	GO:0004252; F:serine-type endopeptidase activity; IEA.			
DR	GO;	GO:0006508; P:proteolysis and peptidolysis; IEA.			
DR	InterPro;	IPR01847; Peptidase S21.			
DR	Pfam;	PF00716; Peptidase S21; I.			
DR	PRINTS;	PR00236; HSCVAPSIDP40.			
KW		Protease.			
SEQUENCE		247 AA; 26619 MW; 0BED160C2DF0AD0C CRC64;			

Query Match 33.4%; Score 1131; DB 2; Length 247;

Best Local Similarity 90.3%; Pred. No. 7e-44;

Matches 223; Conservative 8; Mismatches 16; Indels 0; Gaps 0;

Qy	1	MASAEMLERLEAPLPDRAVPIYVAGFLALYDSGDPGELALDPDTVRAALPPENPLPIND	60
Db	1	MAADAPGDREMEELPDRAPVIYVAGFLALYDSGSELALDPDTVRAALPDENPLPIND	60
Qy	61	HRARCEGRVLAVVNDPRGPPFVGLIACVQLERVLETAASAAIIFERRGPALSRERLIYL	120

Db 61 HRACGVRVLA VDDPRGPFVFFGLIACVQLERVLETAASAAI FERRGGPIISRERLLYL 120

Qy 121 ITNYLPSVLSLTKRRGDEVPDRTLFAHVALCAIGRLGTITVTYDTSLDAAIAPFRHLDP 180

Db 121 ITNYLPSVLSLTKRLGGEAHPDRTLFAHVALCAIGRLGTITVTYDGLDAAIAPFRHLSP 180

Qy 181 ATREGVREAAEAELALAGRTWAPGVBALTHTLTSTAVNNMMLDRWSLVAERRRQAGIA 240

Db 181 ASREGARLAAEAELASLGRTPAGVEALHTLLTSTAVNNMMLDRWSLVAERRRQAGIA 240

Qy 241 GHYLYQA 247

Db 241 GHYLYQA 247

RESULT 11

VP40_EHV1B STANDARD; PRT; 646 AA.

AC P28936; Q69263;

DT 01-DEC-1992 (Rel. 24, Created)

DT 01-DEC-1992 (Rel. 24, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Capsid protein P40 (Varion structural gene 35 protein) [Contains:

DE Capsid protein VP24 (Assemblin) (Protease) (EC 3.4.21.97); Capsid

DE protein VP22A).

DE Names=35;

GN Equine herpesvirus 1 (strain Ab4p) (EHV-1) (Equine abortion virus).

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

OC Alphaherpesvirinae; Varicellovirus.

OX NCBI_TaxID=31520;

OX 11

RP SEQUENCE FROM N.A.

RN MEDLINE=92295566; PubMed=1318606;

RX Telford E.A.R., Watson M.S., McBride K., Davison A.J.;

RA "The DNA sequence of equine herpesvirus-1.";

RT Virology 189:304-316(1992).

RL -1- FUNCTION: VP22A is a component of the capsid core involved in processing and packaging of progeny DNA. VP24 is a protease which can proteolytically cleave itself and VP22A at the C-terminus.

CC -1- CATALYTIC ACTIVITY: Cleaves -Ala-|-Ser- and -Ala-|-Ala- bonds in the scaffold protein.

CC -1- SIMILARITY: Belongs to peptidase family S21.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.ebi.ac.uk/ebis-sib.ch/announcement> or send an email to license@ebi-sib.ch).

CC EMBL; M86664; AAB02470.1; --

DR EMBL; M86664; AAB02471.1; --

DR PIR; I36798; WZBEC8.

DR HSP; Q69527; IAT3.

DR MEROPS; S21.001; --

DR InterPro; IPR001847; Peptidase S21.

DR Pfam; PF00716; Peptidase_S21; 1.

DR PRINTS; PR00236; HSCAPSIDP40.

FW Capsid assembly; Coat protein; Hydrolyase; Serine protease.

KT CHAIN 1 646

FT CHAIN 318 646

FT CHAIN 1 242

FT CHAIN 243 7622

FT PROPEP 7623 646

FT SITE 242 243

FT SITE 622 623

FT ACT_SITE 55 55

FT ACT_SITE 123 123

FT ACT_SITE 142 142

SEQUENCE 646 AA; 68579 MW; 1C1C28C848116276 CRC64;

Query Match 31.9%; Score 1079.5; DB 1; Length 646;

```
Best Local Similarity 40.3%; Pred. No. 3.6e-41;
Matches 278; Conservative 75; Mismatches 212; Indels 125; Gaps 20;

QY 18 AVPIYVAGFLALYDSDGDELALDDPTVRAALPPENPLPINVDHRRARCEVGRVLAVNDP 77
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
12 SLPIYVAGYIALYDMGCGELTLTRETVAALPPASRLPINIDHRNGCVGVLSTVDDA 71

QY 78 RGPFFVGLIACVQLERVLTAASAAIFERRGPAISREERLLYLTNYLPSVLSLTKRGD 137
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
72 RGPFFLGIINCPLQAGVLAATAGPFFGELSGELSEQRLLYLVSNYLPASLSRRLGP 131

QY 138 EYPPDRTLFAHVALCAIGRRLLGTIVYDTSLSAAIAPFRHLDPATREGVRRRAAEALAL 197
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
132 DEPEDTLFAHVSCLVIGRRVGTIVYDATPENAVAPFKRLSPSSRELLITAREAQSL 191

QY 198 A-GRTWAPGVEALHTLLSTAVNNMLDRSLVAERROAGIAGHTYLOASEKFKI--- 253
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
192 GDAATWHLSEDTLTRVLLSTAVNNMLLRNWLVARRRREAGIEGHTYLOASASFGITNG 251

QY 254 -----WGAESAPAPERGVKGTGAPGAMDTSPAASVPAQVAVRAROVASS----- 298
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
252 CNKADFCGAELVDT--CGYKSG-----EKVHGAPYSRVTLGAKAFTSSSPNALPSSDN 302

QY 299 -----SSSGFPAPADMNVPVSAGAPAPPPGDSGLYMWIPASHYNQLVTGQSAPRHP 350
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
303 DKGGIGERTQKHISAWASSNPQTLNAGAPLVSGD--YILVPAQYNQLVVGQHT-SHPP 359

QY 351 LTAAGLPAAGTVAYGHGPGAGSPHYPPPAHPYPMGLFAGP-----SPLAQIAALV 402
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
360 INAGPAPVTHAV-----PSQYTPPAYNSLMPSPMYQAPPVSWPHSANLEAQITALV 411

QY 403 GAIAADRCAGGLPAAAGDHGIRGS-----AKRRRHEVEQPEYDCGRDEPDR 448
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
412 GALAADRKA---TKGSDPHVIQGSQCSPLSPQQRERYARKRHWDATTRD-----DL 462

QY 449 DFPYYPGEARPEPRVDSRRARQASGPHETITLAVGAVTSLQQLAHMRA-----RT 501
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
463 EGIYYPGERSPRP---GERRAGR-----PSTTIADLMGAVSSLQOEVSQALRQTVAQPO 515

QY 502 HAPYGPYPVGVYHHPHADTETPAQPPRYPAKAVLPPPH-----IAPGPPLSGAVPPP 556
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
516 AAPAGLYKPI-----PAVPPQY-SQYQIQPOHVAISAVAPOLPGI-----P 556

QY 557 SYPPVAVTFGPAPPLHQPSPAHAHPPPPPGTTPPAASLPOPEAPGAAGALVNASSAA 616
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
557 SQFTQAVLAPQVPAAGEAGSAGKVAASAPQAQAEARAPQOFEAVTSAVLVPTQPOQAS 616

QY 617 HVNVDTARA-----ADLFVSQMMGSR 637
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
617 SQTVDASASTGLEFGRDDADIFVSQMMGAR 646

RESULT 12
Q6S6T7 PRELIMINARY; PRT; 646 AA.
AC Q6S6T7
DT 05-JUL-2004 (trEMBLrel. 27, Created)
DT 05-JUL-2004 (trEMBLrel. 27, Last sequence update)
DT 01-OCT-2004 (trEMBLrel. 28, Last annotation update)
DE Capsid protein (Protease).
OS Equine herpesvirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10326;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=V592;
RA Davis-Poynter N.J., Nugent J., Birch-Machin I., Allen G.P.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=Ab4;
RX MEDLINE=92295566; PubMed=1318606;
```

RESULT 14
Q6X238
ID Q6X238

RESULT 15

AA86135 PRELIMINARY; PRT; 619 AA.
AC AAR86135;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE UL26 capsid maturation serine protease.
GN BHV5-31.
OS Bovine herpesvirus 5.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae.
OX NCBI_TaxID=35244;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SV507/99;
RX MEDLINE=22850801; PubMed=12970418;
RA Delhon G., Moraes M.P., Lu Z., Afonso C.L., Flores E.F., Weiblen R.,
RA Kutish G.F., Rock D.L.;
RT "Genome of bovine herpesvirus 5."
RL J. Virol. 77:10339-10347(2003).
DR EMBL; AY261359; AAR86135.1; -.
KW Protease.
SQ SEQUENCE 619 AA; 63656 MW; 19275E8F26930B84 CRC64;

Query Match 31.2%; Score 1056; DB 2; Length 619;
Best Local Similarity 41.4%; Pred. NO. 4e-40;
Matches 270; Conservative 63; Mismatches 229; Indels 90; Gaps 19;
QY 18 AVPIYVAGFLAYDSDGDELDPDPTVRAALPPENPLPINVDHRARCEVGRVLAVVNDP 77
TQ 26 SMPVYVGGYALYGMDEGELVLTREVARALPPAAPLPINIDHASACEVGAVLALADDD 85
QY 78 RGPFFVGLIACVQLERVLTAASAAIFERRRGALSRERLLYLITNPLSPVSLSTKRRGD 137
DB 86 AGLFFVGVVNCQLADVLADVAHPAFFGADAPALAPRERFLVLVSNLPSVLSRRLAP 145
QY 138 EYPPDRTLFAHVALCAIGRLGTIVTYDTSLSDAATAPFRHLDPATREGVGRRAAEALAL 197
DB 146 GSEADGTLFAHVALCVLGRVGTIVTYDATPEACVAPFRRLSPRARAALLADAEEAARAL 205
QY 198 AGRTWAPGVREALTHLLSTAVNNMLRDRSLVAERROAGIAGHTYLOASEKFKIWGAE 257
DB 206 GDRAMPVPREALARTLLSTAVNNMLVRDKWDTVSRRRREAGIAGHTYLOASAVFFPLAGGE 265
QY 258 SAPAPERGVKTGAPGAMDTSPAASVPAPQVAVRQVAVSSSSSSSPAPADMNVPVSASGA 317
DB 266 E-PERSAQKAGLAGV-----CIALPVASGEARQPELPPA-----PPPMASAVHQASAS 314
QY 318 PAPPFPGDGSYLWIPASHYNQLVTQS-----APRH-PPLTAC-----GL 356
DB 315 PAHFLPA-GDYVVVPTAQYNQLVVSQARGAAAAAPPAPVFLPPALAAAPPMPMGWYGA 373
QY 357 PAAGTVAYGHPGAGSPHPPPPAHPYPGMLFAGSPLEAQIAALVGAIAADR--QAGGL 414
DB 374 AASGAAAPWHPGYG-----FPPP-----GLESQIMALAGATDGRRLQAQGA 415
QY 415 --PAAAGDHGIRGSAKRREHEVQPEYDCGRDEPDPDFYPYGEARPEPRPVDSRRAARQ 472
DB 416 DGGYDGLDRPPAKRRYNNWPLESRGGD-----DEAYYPGEGAPAEPLPQHHRPPSP 471
QY 473 ASGPETHITVALGAVTSLQOELAHMRARTHAPYGPYPVGVYHHPHADTETPAQPPRYPA 532
DB 472 PQAPHALSRLASAVSSSLQQEVSQRLA--GYPYGPAFAAAQHLP-----PAPVCLPQ 523
QY 533 KAVILPPPHIAPGPPPLSGAVPPPSY---PPVAVTEGPAPPLHQPSPAHAHPPPPPPGPT 589
DB 524 QQ-QPPQQAAPPQPVVAGPAPAPQAALAPAPAQAAPAPPLATPA-----A 569
QY 590 PPAASLPOPEAPGAEGALVNASSAAHVNV-----DTARAADLFVQMMGSR 637

Db 570 PAPAAAGGPPEPGAI--ATVDASAVAGLPLAQPFQACDPADIFVAQMMRHR 619

Search completed: January 10, 2005, 19:18:33
Job time : 210 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	2333	68.9	635	4	US-08-176-320-2	Sequence 2, Appli
2	2323	68.6	635	1	US-07-832-855-2	Sequence 2, Appli
3	1432	42.3	350	1	US-08-458-067-2	Sequence 2, Appli
4	1432	42.3	350	5	PCR-US96-07795-2	Sequence 2, Appli
5	1432	42.3	350	5	PCR-US96-07796-2	Sequence 2, Appli
6	1262	37.3	306	1	US-08-279-754-2	Sequence 2, Appli
7	1262	37.3	306	5	PCR-US95-09052-2	Sequence 2, Appli
8	1252	37.0	247	3	US-09-167-434-4	Sequence 4, Appli
9	1252	37.0	247	3	US-08-853-755-4	Sequence 4, Appli
10	1126	33.3	247	3	US-09-167-434-3	Sequence 3, Appli
11	1126	33.3	247	3	US-08-853-755-3	Sequence 3, Appli
12	787	23.2	663	1	US-07-912-015-2	Sequence 2, Appli
13	590.5	17.4	243	3	US-09-167-434-9	Sequence 9, Appli
14	590.5	17.4	243	3	US-08-853-755-9	Sequence 9, Appli
15	590.5	17.4	255	3	US-09-167-434-10	Sequence 10, Appl
16	590.5	17.4	255	3	US-08-853-755-10	Sequence 10, Appl
17	586.5	17.3	236	3	US-09-167-434-5	Sequence 5, Appli
18	586.5	17.3	236	3	US-08-853-755-5	Sequence 5, Appli
19	586.5	17.3	242	3	US-09-167-434-8	Sequence 8, Appli
20	586.5	17.3	242	3	US-08-853-755-8	Sequence 8, Appli
21	586.5	17.3	246	3	US-09-167-434-11	Sequence 11, Appl
22	586.5	17.3	246	3	US-08-853-755-11	Sequence 11, Appl
23	586.5	17.3	254	3	US-09-167-434-7	Sequence 7, Appli
24	586.5	17.3	254	3	US-08-853-755-7	Sequence 7, Appli
25	329.5	9.7	609	1	US-07-798-776-2	Sequence 2, Appli
26	329.5	9.7	609	3	US-08-251-288A-2	Sequence 2, Appli
27	329.5	9.7	609	3	US-09-298-819A-2	Sequence 2, Appli

```

Db 61 HRAGEVGRVAVDDPRGFFVGLIACVQLSERVLETAASAAIFERRGPPLSREERL120
QY 121 ITNYLPSVSLTKRGRDEVPPDRTLFAHVALCAIGRRLLTIVYDTSDDAAIAPRHLDP 180
Db 121 ITNYLPSVSLATKRGGEAHPDRTLFAHVALCAIGRRLLTIVYDTSDDAAIAPRHLSP 180
QY 181 ATREGVREAAEALAGRTWAPGVEALHTLLSTAVNNMLDRWSLVAERROAGIA 240
Db 181 ASREGARRLAEEALSGRTWAPGVEALHTLLSTAVNNMLDRWSLVAERROAGIA 240
QY 241 GHTYLOASEKFKINGAESAPAPRGYKTCGAPAMDTSPAASVPAPO---VAVRARQVAS 296
Db 241 GHTYLOASEKFKINGAESAPAPRGYKTCGAPAMDTSPAASVPAPO---VAVRARQVAS 300
QY 297 SSSSSFPAPADNVPVSAAGAPAPPPGSGSYLWIPASHYNQLVTGQSPAPRHPPLTACGL 356
Db 301 S-----PVLPPMNPVPTSGTPAPAPPGDGSYLWIPASHYNQLVAGHAAPQOPHSAFGF 354
QY 357 P-AAGTVAGHPGAGSPHYPPPAHPYPCMLFAGSPLEAQIAALVGNIAADROAGGLP 415
Db 355 PAAAGSVAYGHPGAGLSQHQELAHMRARTSAPYGMVTPVAHYRPOVGEPEPTTTHPALCPPEA 414
QY 416 AAAGDHGIRGSAKRRRHEVEQPEYDCGRDEPRDPPYYPGEARPEPRPVDSRRAARQASG 475
Db 415 -AAGDPGVGRSGKRRRYEAGPSESICDQDEPDADYYPYFGEARGAPRGVDSRRAARHSPG 473
QY 476 PHETITLVGAVTSLQOELAHMRARTHAPYGPYPVGYHHPHADTE--TPAQPPYPYPAKA 534
Db 474 TNETITLMGAVTSLQOELAHMRARTSAPYGMVTPVAHYRPOVGEPEPTTTHPALCPPEA 533
QY 535 VYLPPLPHIAPGPPPLSGA--VPPPSYPVAVTPGAPPLHQSPAHAPPPPPPGTPTPP 592
Db 534 VYRPPHSAFYGPPOGPASHAPTTPYAPAACTPPGPPPP---PCPSTQTRAPLPTTEPAFP 590
QY 593 AASLPQEPAPGAEGALVNASAAHVNVDTRAADLFVSQMMGR 637
Db 591 AATGQPEASNAEAGALVNASAAHVDVDTARAADLFVSQMMGAR 635

```

RESULT 2

```

US-07-832-855-2
; Sequence 2, Application US/07832855
; Patent No. 5478727
; GENERAL INFORMATION:
; APPLICANT: Roizman, Bernard
; APPLICANT: Liu, FenYong
; TITLE OF INVENTION: Methods and Compositions of a
; TITLE OF INVENTION: Preparation and Use of A Herpes Protease
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARNOLD, WHITE & DURKEE
; STREET: 321 No. 5478727th Clark Street, Suite 800
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60610
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/832,855
; FILING DATE: 19920207
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coolley, Ronald B.
; REGISTRATION NUMBER: 27,187
; REFERENCE/DOCKET NUMBER: ARCD045
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 744-0090
; TELEFAX: (312) 245-4961

```

```

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 635 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-832-855-2

```

```

Query Match 68.6%; Score 2323; DB 1; Length 635;
Best Local Similarity 73.6%; Pred. No. 1e-159;
Matches 475; Conservative 30; Mismatches 122; Indels 18; Gaps 7;

```

```

QY 1 MASAEEMRRLRAPLDPRAVPIYVAGFLALYDSDGDELALDPDVTVAALPPNPPIINV 60
Db 1 MAADAPGRMEPELPDRAVPIYVAGFLALYDSDGDELALDPDVTVAALPPNPPIINV 60
QY 61 HRACEVGRVAVNDPRGPPFVGLIACVQLSERVLETAASAAIFERRGPPLSREERL120
Db 61 HRACEVGRVAVNDPRGPPFVGLIACVQLSERVLETAASAAIFERRGPPLSREERL120
QY 121 ITNYLPSVSLTKRGRDEVPPDRTLFAHVALCAIGRRLLTIVYDTSDDAAIAPRHLDP 180
Db 121 ITNYLPSVSLATKRGGEAHPDRTLFAHVALCAIGRRLLTIVYDTSDDAAIAPRHLSP 180
QY 181 ATREGVREAAEALAGRTWAPGVEALHTLLSTAVNNMLDRWSLVAERROAGIA 240
Db 181 ASREGARRLAEEALSGRTWAPGVEALHTLLSTAVNNMLDRWSLVAERROAGIA 240
QY 241 GHTYLOASEKFKINGAESAPAPRGYKTCGAPAMDTSPAASVPAPO---VAVRARQVAS 296
Db 241 GHTYLOASEKFKINGAESAPAPRGYKTCGAPAMDTSPAASVPAPO---VAVRARQVAS 300
QY 297 SSSSSFPAPADNVPVSAAGAPAPPPGSGSYLWIPASHYNQLVTGQSPAPRHPPLTACGL 356
Db 301 S-----PVLPPMNPVPTSGTPAPAPPGDGSYLWIPASHYNQLVAGHAAPQOPHSAFGF 354
QY 357 P-AAGTVAGHPGAGSPHYPPPAHPYPCMLFAGSPLEAQIAALVGNIAADROAGGLP 415
Db 355 PAAAGSVAYGHPGAGLSQHQELAHMRARTHAPYGPYPVGYHHPHADTE--TPAQPPYPYPAKA 473
QY 416 AAAGDHGIRGSAKRRRHEVEQPEYDCGRDEPRDPPYYPGEARPEPRPVDSRRAARQASG 475
Db 415 -AAGDPGVGRSGKRRRYEAGPSESICDQDEPDADYYPYFGEARGAPRGVDSRRAARHSPG 473
QY 476 PHETITLVGAVTSLQOELAHMRARTHAPYGPYPVGYHHPHADTE--TPAQPPYPYPAKA 534
Db 474 TNETITLMGAVTSLQOELAHMRARTSAPYGMVTPVAHYRPOVGEPEPTTTHPALCPPEA 533
QY 535 VYLPPLPHIAPGPPPLSGA--VPPPSYPVAVTPGAPPLHQSPAHAPPPPPPGTPTPP 592
Db 534 VYRPPHSAFYGPPOGPASHAPTTPYAPAACTPPGPPPP---PCPSTQTRAPLPTTEPAFP 590
QY 593 AASLPQEPAPGAEGALVNASAAHVNVDTRAADLFVSQMMGR 637
Db 591 PPTGQPEASNAEAGALVNASAAHVDVDTARAADLFVSQMMGAR 635

```

RESULT 3

```

US-08-458-067-2
; Sequence 2, Application US/08458067
; Patent No. 5728557
; GENERAL INFORMATION:
; APPLICANT: Register, Robert B.
; APPLICANT: Shafer, Jules A.
; TITLE OF INVENTION: HERPES SIMPLEX TYPE 1 PROTEASE MUTANTS
; TITLE OF INVENTION: AND VECTORS
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ms. Joanne M. Giesser
; STREET: 126 East Lincoln Avenue, P.O. Box 2000-0907
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: US

```


; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458,067
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Giesser, Joanne M.
; REGISTRATION NUMBER: 32,838
; REFERENCE/DOCKET NUMBER: 19457
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-3046
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 350 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-458-067-2

Query Match 42.3%; Score 1432; DB 1; Length 350;
Best Local Similarity 80.5%; Pred. No. 9.9e-96;
Matches 285; Conservative 13; Mismatches 46; Indels 10; Gaps 2;

Qy	1	MASAEERERLEAPLDRAPVIYVAGFLALYDSGDPGELALDPDTVRAALPPENPLPINVD	60
Db	1	MAADAPGDMEELPDRAVPIYVAGFLALYDSGSGELALDPDTVRAALPPDNPLPINVD	60
Qy	61	HRACEVGRVLAVVNDPRGPFVGLIACVOLERVLETAASAAIFERRGPALSREERLLYL	120
Db	61	HRACEVGRVLAVVDDPRGPFVGLIACVOLERVLETAASAAIFERRGPALSREERLLYL	120
Qy	121	ITNYPVSLSTKRGDEVPDRTLFAHVALCAIGRLGTVITYDTSLDAAIAPFRHLDP	180
Db	121	ITNYPVSLATKRLGGEAHPDRTLFAHVALCAIGRLGTVITYDTGLDAAIAPFRHLSP	180
Qy	181	ATREGVREAAEAELALAGRTWAPGVEALHTLLSTAVNNMLDRWSLVAERRRQAGIA	240
Db	181	ASREGARRLAAEAELALSGRTWAPGVEALHTLLSTAVNNMLDRWSLVAERRRQAGIA	240
Qy	241	GHTYLOASEKFKIWAESAAPAPRGYKTAGAMDTSPAASVPAQ-----VAVRAROVAS	296
Db	241	GHTYLOASEKFKIWAESAAPAPRGYKTAGAMDTSPAASVPAQ-----VAVRAROVAS	296
Qy	297	SSSSSPFAPADMPVSVASGAPAPPPGDSYLWIPASHYNQLVTGOSAPRHPP	350
Db	301	P-----PVLPPMNPVPTSGTPAPAPPGDSYLWIPASHYNQLVAGHAAPQPOP	348

RESULT 4
PCT-US96-07795-2
; Sequence 2, Application PC/TUS9607795
; GENERAL INFORMATION:
; APPLICANT: MERCK & CO., INC.
; APPLICANT: Registrar, Robert B.
; APPLICANT: Shafer, Jules A.
; TITLE OF INVENTION: HERPES SIMPLEX TYPE 1 PROTEASE MUTANTS
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ms. Joanne M. Giesser
; STREET: 126 East Lincoln Avenue, P.O. Box 2000-0907
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: US
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/07795
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Giesser, Joanne M.
; REGISTRATION NUMBER: 32,838
; REFERENCE/DOCKET NUMBER: 19457
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-3046
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 350 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US96-07795-2

Query Match 42.3%; Score 1432; DB 5; Length 350;
Best Local Similarity 80.5%; Pred. No. 9.9e-96;
Matches 285; Conservative 13; Mismatches 46; Indels 10; Gaps 2;

Qy	1	MASAEERERLEAPLDRAPVIYVAGFLALYDSGDPGELALDPDTVRAALPPENPLPINVD	60
Db	1	MAADAPGDMEELPDRAVPIYVAGFLALYDSGSGELALDPDTVRAALPPDNPLPINVD	60
Qy	61	HRACEVGRVLAVVNDPRGPFVGLIACVOLERVLETAASAAIFERRGPALSREERLLYL	120
Db	61	HRACEVGRVLAVVDDPRGPFVGLIACVOLERVLETAASAAIFERRGPALSREERLLYL	120
Qy	121	ITNYPVSLSTKRGDEVPDRTLFAHVALCAIGRLGTVITYDTSLDAAIAPFRHLDP	180
Db	121	ITNYPVSLATKRLGGEAHPDRTLFAHVALCAIGRLGTVITYDTGLDAAIAPFRHLSP	180
Qy	181	ATREGVREAAEAELALAGRTWAPGVEALHTLLSTAVNNMLDRWSLVAERRRQAGIA	240
Db	181	ASREGARRLAAEAELALSGRTWAPGVEALHTLLSTAVNNMLDRWSLVAERRRQAGIA	240
Qy	241	GHTYLOASEKFKIWAESAAPAPRGYKTAGAMDTSPAASVPAQ-----VAVRAROVAS	296
Db	241	GHTYLOASEKFKIWAESAAPAPRGYKTAGAMDTSPAASVPAQ-----VAVRAROVAS	296
Qy	297	SSSSSPFAPADMPVSVASGAPAPPPGDSYLWIPASHYNQLVTGOSAPRHPP	350
Db	301	P-----PVLPPMNPVPTSGTPAPAPPGDSYLWIPASHYNQLVAGHAAPQPOP	348

RESULT 5
PCT-US96-07796-2
; Sequence 2, Application PC/TUS9607796
; GENERAL INFORMATION:
; APPLICANT: MERCK & CO., INC.
; APPLICANT: Registrar, Robert B.
; APPLICANT: Shafer, Jules A.
; TITLE OF INVENTION: HERPES SIMPLEX TYPE 1 PROTEASE MUTANTS
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ms. Joanne M. Giesser
; STREET: 126 East Lincoln Avenue, P.O. Box 2000-0907
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: US
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30


```

; TELEPHONE: (908) 594-3046
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 306 amino acids
; TYPE: amino acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
PCT-US95-09052-2

Query Match 37.3%; Score 1262; DB 5; Length 306;
Best Local Similarity 84.0%; Pred. No. 1.6e-83;
Matches 252; Conservative 11; Mismatches 33; Indels 4; Gaps 1;

QY 1 MASAENRERLEAPLPDRAPIYVAGFLALYDSDGPGELALDPTVRAALPPENPLPINVD 60
Db 1 MAADAFQDRMEELPDRAPIYVAGFLALYDSDGPGELALDPTVRAALPPENPLPINVD 60
QY 61 HRARCEVGRVLAIVNDPRGPFVGLIACVQLERVLETAASAAIFERRGPALSREERLLYL 120
Db 61 HRARCEVGRVLAIVNDPRGPFVGLIACVQLERVLETAASAAIFERRGPALSREERLLYL 120
QY 121 ITNLPVSLSLTKRGDEVPDPTLFAHVALCAIGRRLGITVYDTSLDAAIAPFRHLDP 180
Db 121 ITNLPVSLSLTKRGDEVPDPTLFAHVALCAIGRRLGITVYDTSLDAAIAPFRHLSP 180
QY 181 ATREGVRRREAAEALAGRTWAPGVLEALHTLLSTAVNNMLRDRWSLVAERRRQAGIA 240
Db 181 ASREGVRRREAAEALAGRTWAPGVLEALHTLLSTAVNNMLRDRWSLVAERRRQAGIA 240
QY 241 GHTYLOASEKFKTWGAESAPAPRGYKGTAPGAMDTSPAASVPAPQ---VAVRARQVAS 296
Db 241 GHTYLOASEKFKTWGAESAPAPRGYKGTAPGAMDTSPAASVPAPQ---VAVRARQVAS 300

RESULT 8
US-09-167-434-4
; Sequence 4, Application US/09167434
; Patent No. 6008033
; GENERAL INFORMATION:
; APPLICANT: Abdel-Meguid, Sherin
; APPLICANT: Qiu, Xiayang
; APPLICANT: Culp, Jeffrey
; APPLICANT: Debouck, Christine
; APPLICANT: Janson, Cheryl A.
; APPLICANT: Hoog, Susan S.
; APPLICANT: Smith, Ward W.
; TITLE OF INVENTION: No. 6008033el Proteases, Compositions Capable of
; Binding to Said Site, and Methods of Use Thereof
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; STREET: 709 Swedeland Road - P.O. Box 1539
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19406-2799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/167,434
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/853,755
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/022,470

```

```

; FILING DATE: 26-JUL-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,416
; FILING DATE: 21-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/030,901
; FILING DATE: 14-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/035,973
; FILING DATE: 21-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/039,191
; FILING DATE: 27-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Dinner, Dara L.
; REGISTRATION NUMBER: 33,680
; REFERENCE/DOCKET NUMBER: P50472-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5022
; TELEFAX: 610-270-5090
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 247 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-167-434-4

Query Match 37.0%; Score 1252; DB 3; Length 247;
Best Local Similarity 99.6%; Pred. No. 6.4e-83;
Matches 246; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASAENRERLEAPLPDRAPIYVAGFLALYDSDGPGELALDPTVRAALPPENPLPINVD 60
Db 1 MASAENRERLEAPLPDRAPIYVAGFLALYDSDGPGELALDPTVRAALPPENPLPINV 60
QY 61 HRARCEVGRVLAIVNDPRGPFVGLIACVQLERVLETAASAAIFERRGPALSREERLLYL 120
Db 61 HRARCEVGRVLAIVNDPRGPFVGLIACVQLERVLETAASAAIFERRGPALSREERLLYL 120
QY 121 ITNLPVSLSLTKRGDEVPDPTLFAHVALCAIGRRLGITVYDTSLDAAIAPFRHLDP 180
Db 121 ITNLPVSLSLTKRGDEVPDPTLFAHVALCAIGRRLGITVYDTSLDAAIAPFRHLDP 180
QY 181 ATREGVRRREAAEALAGRTWAPGVLEALHTLLSTAVNNMLRDRWSLVAERRRQAGIA 240
Db 181 ATREGVRRREAAEALAGRTWAPGVLEALHTLLSTAVNNMLRDRWSLVAERRRQAGIA 240
QY 241 GHTYLOA 247
Db 241 GHTYLOA 247

RESULT 9
US-08-853-755-4
; Sequence 4, Application US/08853755
; Patent No. 6083711
; GENERAL INFORMATION:
; APPLICANT: Abdel-Meguid, Sherin
; APPLICANT: Qiu, Xiayang
; APPLICANT: Culp, Jeffrey
; APPLICANT: Debouck, Christine
; APPLICANT: Janson, Cheryl A.
; APPLICANT: Hoog, Susan S.
; APPLICANT: Smith, Ward W.
; TITLE OF INVENTION: No. 6083711el Proteases, Compositions Capable of
; Binding to Said Site, and Methods of Use Thereof
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; STREET: 709 Swedeland Road - P.O. Box 1539
; CITY: King of Prussia

```

STATE: Pennsylvania
COUNTRY: USA
ZIP: 19406-2799
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICANT: Hoog, Susan S.
FILING DATE: 21-AUG-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/018,616
FILING DATE: 15-MAY-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/022,470
FILING DATE: 26-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/024,416
FILING DATE: 21-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/030,901
FILING DATE: 14-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/035,973
FILING DATE: 21-JAN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/039,191
FILING DATE: 27-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Dinner, Dara L.
REGISTRATION NUMBER: 33,680
REFERENCE/DOCKET NUMBER: P50472-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5022
TELEFAX: 610-270-5090
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 247 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-853-755-4

Query Match 37.0%; Score 1252; DB 3; Length 247;
Best Local Similarity 99.6%; Pred. No. 6.4e-83;
Matches 246; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASAEMRERLEAPLPDRAVPIYVAGFLALYDSGDPGELALDPTVRAALPPENPLPINDV 60
DB 1 MASAEMRERLEAPLPDRAVPIYVAGFLALYDSGDPGELALDPTVRAALPPENPLPINV 60
QY 61 HRAECVGRVLA VNDPRGPFVGLIACVQLSERVLETAASAIFERRGFPALSRERLLYL 120
DB 61 HRAECVGRVLA VNDPRGPFVGLIACVQLSERVLETAASAIFERRGFPALSRERLLYL 120
QY 121 ITNYLPSVSLTKRGDEVPDRTFAHVALCAIGRLGTIVTYDTSLDAAIAPFRHLDP 180
DB 121 ITNYLPSVSLTKRGDEVPDRTFAHVALCAIGRLGTIVTYDTSLDAAIAPFRHLDP 180
QY 181 ATREGVRRRAAEALAGRTWAGVEALHTLLSTAVNNMLRDRWSLVAERRRQAGIA 240
DB 181 ATREGVRRRAAEALAGRTWAGVEALHTLLSTAVNNMLRDRWSLVAERRRQAGIA 240
QY 241 GHTYLOA 247
DB 241 GHTYLOA 247

RESULT 10

US-09-167-434-3

Sequence 3, Application US/09167434
Patent No. 6008033
GENERAL INFORMATION:
APPLICANT: Abdel-Meguid, Sherin
APPLICANT: Qiu, Xiyang
APPLICANT: Culp, Jeffrey
APPLICANT: Debouck, Christine
APPLICANT: Janson, Cheryl A.
APPLICANT: Hoog, Susan S.
APPLICANT: Smith, Ward W.
TITLE OF INVENTION: No. 6008033el Proteases, Compositions Capable of
TITLE OF INVENTION: Binding to Said Site, and Methods of Use Thereof
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road - P.O. Box 1539
CITY: King of Prussia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19406-2799
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/167,434
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/853,755
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/022,470
FILING DATE: 26-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/024,416
FILING DATE: 21-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/030,901
FILING DATE: 14-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/035,973
FILING DATE: 21-JAN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/039,191
FILING DATE: 27-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Dinner, Dara L.
REGISTRATION NUMBER: 33,680
REFERENCE/DOCKET NUMBER: P50472-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5022
TELEFAX: 610-270-5090
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 247 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-167-434-3

Query Match 33.3%; Score 1126; DB 3; Length 247;

Best Local Similarity 89.9%; Pred. No. 8e-74;

Matches 222; Conservative 9; Mismatches 16; Indels 0; Gaps 0;

QY 1 MASAEMRERLEAPLPDRAVPIYVAGFLALYDSGDPGELALDPTVRAALPPENPLPINDV 60
DB 1 MAAAPAGDRMEEPDRAVPIYVAGFLALYDSGDPGELALDPTVRAALPPDNPLPINV 60
QY 61 HRAECVGRVLA VNDPRGPFVGLIACVQLSERVLETAASAIFERRGFPALSRERLLYL 120

```

Db      61 HRACCEVGRVLAVVDDPRGPFVGLIACVQLERVLETAASAAIFERRGPPLSREERLLYL 120
Qy      121 ITNLPVSLSLTKRGGDEVPDRTLFAHVALCAIGRLGTVITVDTSLDAAIAPFRHLDP 180
Db      121 ITNLPVSLSLTKRGGDEVPDRTLFAHVALCAIGRLGTVITVDTSLDAAIAPFRHLSP 180
Qy      181 ATREGVRRERAAEALAGRTWAPGVEALHTLLSTAVNNMLRDRWSLVAERRRQAGIA 240
Db      181 ASREGARRLAEEAELALSGRTWAPGVEALHTLLSTAVNNMLRDRWSLVAERRRQAGIA 240
Qy      241 GHTYLOA 247
Db      241 GHTYLOA 247

```

```

RESULT 11
US-08-853-755-3
; Sequence 3, Application US/08853755
; Patent No. 6083711
; GENERAL INFORMATION:
; APPLICANT: Abdel-Meguid, Sherin
; APPLICANT: Qiu, Xiayang
; APPLICANT: Culp, Jeffrey
; APPLICANT: Debouck, Christine
; APPLICANT: Janson, Cheryl A.
; APPLICANT: Hoog, Susan S.
; APPLICANT: Smith, Ward W.
; TITLE OF INVENTION: No. 6083711el Proteases, Compositions Capable of
; BINDING TO SAID SITE, AND METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road - P.O. Box 1539
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19406-2799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/853,755
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/018,616
; FILING DATE: 15-MAY-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/022,470
; FILING DATE: 26-JUL-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,416
; FILING DATE: 21-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/030,901
; FILING DATE: 14-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/035,973
; FILING DATE: 21-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/039,191
; FILING DATE: 27-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Dinner, Dara L.
; REGISTRATION NUMBER: 33,680
; REFERENCE/DOCKET NUMBER: P50472-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5022
; TELEFAX: 610-270-5090
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 247 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-853-755-3
; Query Match 33.3%; Score 1126; DB 3; Length 247;
; Best Local Similarity 89.9%; Pred. No. 8e-74;
; Matches 222; Conservative 9; Mismatches 16; Indels 0; Gaps 0;
Qy      1 MASAEMRERLEAPLPPRAVPIYVAGFLALYDSGDPGELALDPDTVRAALPPENPLPINVD 60
Db      1 MAADAPGDRMEELPDRAPVIYVAGFLALYDSGDPGELALDPDTVRAALPPENPLPINVN 60
Qy      61 HRACCEVGRVLAVVNDPRGPFVGLIACVQLERVLETAASAAIFERRGPPLSREERLLYL 120
Db      61 HRACCEVGRVLAVVNDPRGPFVGLIACVQLERVLETAASAAIFERRGPPLSREERLLYL 120
Qy      121 ITNLPVSLSLTKRGGDEVPDRTLFAHVALCAIGRLGTVITVDTSLDAAIAPFRHLDP 180
Db      121 ITNLPVSLSLTKRGGDEVPDRTLFAHVALCAIGRLGTVITVDTSLDAAIAPFRHLSP 180
Qy      181 ATREGVRRERAAEALAGRTWAPGVEALHTLLSTAVNNMLRDRWSLVAERRRQAGIA 240
Db      181 ASREGARRLAEEAELALSGRTWAPGVEALHTLLSTAVNNMLRDRWSLVAERRRQAGIA 240
Qy      241 GHTYLOA 247
Db      241 GHTYLOA 247
; RESULT 12
; Sequence 2, Application US/07912015
; Patent No. 5283191
; GENERAL INFORMATION:
; APPLICANT: Morgan, Robin Wilson
; APPLICANT: Willemse, Martha Jacoba
; APPLICANT: Claessens, Johannes Antonius Joseph
; APPLICANT: Sondermeijer, Paulus Jacobus Antonius
; TITLE OF INVENTION: Marek's Disease Virus vaccine
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Organon Teknika Corporation
; ADDRESSEE: Biotechnology Research Institute
; STREET: 1330-A Piccard Drive
; CITY: Rockville
; STATE: Maryland
; COUNTRY: U.S.A.
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/912,015
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/615,211
; FILING DATE: 16-NOV-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Donna Bobrowicz
; REGISTRATION NUMBER: 32,196
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301)258-5200
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 663 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

```


Db 230 GHVYLOAS 237

RESULT 14

US-08-853-755-9

; Sequence 9, Application US/08853755

; Patent No. 6083711

; GENERAL INFORMATION:

; APPLICANT: Abdel-Meguid, Sherin

; APPLICANT: Qiu, Xiayang

; APPLICANT: Culp, Jeffrey

; APPLICANT: Debouck, Christine

; APPLICANT: Janson, Cheryl A.

; APPLICANT: Hoog, Susan S.

; APPLICANT: Smith, Ward W.

; TITLE OF INVENTION: No. 6083711el Proteases, Compositions Capable of

; TITLE OF INVENTION: Binding to Said Site, and Methods of Use Thereof

; NUMBER OF SEQUENCES: 18

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SmithKline Beecham Corporation

; STREET: 709 Swedeland Road - P.O. Box 1539

; CITY: King of Prussia

; STATE: Pennsylvania

; COUNTRY: USA

; ZIP: 19406-2799

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/853,755

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/018,616

; FILING DATE: 15-MAY-1996

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/022,470

; FILING DATE: 26-JUL-1996

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/024,416

; FILING DATE: 21-AUG-1996

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/030,901

; FILING DATE: 14-NOV-1996

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/035,973

; FILING DATE: 21-JAN-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/039,191

; FILING DATE: 27-FEB-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Dinner, Dara L.

; REGISTRATION NUMBER: 33,680

; REFERENCE/DOCKET NUMBER: P50472-1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 610-270-5022

; TELEFAX: 610-270-5090

; INFORMATION FOR SEQ ID NO: 9:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 243 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-853-755-9

Query Match 17.4%; Score 590.5; DB 3; Length 243;

Best Local Similarity 50.0%; Pred. No. 3.5e-35;

Matches 124; Conservative 34; Mismatches 79; Indels 11; Gaps 3;

QY 1 MASAEMRERLEAPLPDRAVPIYVAGFIALYDSGDPGLALDDPTVRAALPENPLINVD 60

Db 1 MAAEADECEA-----LYVAGLYALY-SKDEGELNITPEIVRSALPPTSIPINID 51

QY 61 HRARCEVGRVLAVNDPRGPFVGLIACVQLERVIETAASNAIFERRGPAISREERLLYL 120

Db 52 HRKDCVVGEVIAIIEDIRGPFILGIVRCQQLHVLFEAAHSNFFGNRDSVLSPLERALYL 111

QY 121 ITNLPVSVLSLTKRRGDEVPPDRTLFAHVALCAIGRLGTIVTYDTSLDAAIAPFRHLD 180

Db 112 VTNYLPSVLSLSSKRLSNEIPDGNFFTHVALCVGRGVGVVYDCTPESIEFVRVLSM 171

QY 181 ATREGVVRRAEAELALAGTWAPGVREALTHLLSTAVNNMLRDRMSLVAERRRQAGIA 240

Db 172 ESK--ARLLSLVKDYAGLNKVKVSEDKLAQVLLSTAVNNMLLRDRMVDVAKRRREAGIM 229

QY 241 GHVYLOAS 248

Db 230 GHVYLOAS 237

RESULT 15

US-09-167-434-10

; Sequence 10, Application US/09167434

; Patent No. 6008033

; GENERAL INFORMATION:

; APPLICANT: Abdel-Meguid, Sherin

; APPLICANT: Qiu, Xiayang

; APPLICANT: Culp, Jeffrey

; APPLICANT: Debouck, Christine

; APPLICANT: Janson, Cheryl A.

; APPLICANT: Hoog, Susan S.

; APPLICANT: Smith, Ward W.

; TITLE OF INVENTION: No. 6008033el Proteases, Compositions Capable of

; TITLE OF INVENTION: Binding to Said Site, and Methods of Use Thereof

; NUMBER OF SEQUENCES: 18

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SmithKline Beecham Corporation

; STREET: 709 Swedeland Road - P.O. Box 1539

; CITY: King of Prussia

; STATE: Pennsylvania

; COUNTRY: USA

; ZIP: 19406-2799

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/167,434

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/853,755

; FILING DATE:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/022,470

; FILING DATE: 26-JUL-1996

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/024,416

; FILING DATE: 21-AUG-1996

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/030,901

; FILING DATE: 14-NOV-1996

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/035,973

; FILING DATE: 21-JAN-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/039,191

; FILING DATE: 27-FEB-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Dinner, Dara L.

; REGISTRATION NUMBER: 33,680

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2005, 19:18:39 ; Search time 148 Seconds
(without alignments)
1551.918 Million cell updates/sec

Title: US-10-623-429-9

Perfect score: 3385

Sequence: 1 MASAEMRERLEAPLPDRAVP.....VNVDARADLFVQMMSR 637

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1603904 seqs, 360571292 residues

Total number of hits satisfying chosen parameters: 1603904

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/PCT NEW PUB.pdb.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT NEW PUB.pdb.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW PUB.pdb.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pdb.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW PUB.pdb.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pdb.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW PUB.pdb.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pdb.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pdb.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pdb.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pdb.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW PUB.pdb.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pdb.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pdb.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pdb.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pdb.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW PUB.pdb.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW PUB.pdb.*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW PUB.pdb.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pdb.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3385	100.0	637	15	US-10-623-429-9
2	2333	68.9	635	14	Sequence 9, Appli
3	351	10.4	536	17	Sequence 86, Appl
4	294.5	8.7	534	16	Sequence 30, Appl
5	292	8.6	639	14	Sequence 159995
6	284	8.4	825	16	Sequence 2285, Ap
7	281.5	8.3	483	16	Sequence 191498
8	279.5	8.3	569	16	Sequence 149276
9	275.5	8.1	426	16	Sequence 142229
10	273	8.1	506	16	Sequence 126337
11	271.5	8.0	521	16	Sequence 179715
12	270	8.0	606	16	Sequence 195106
13	268.5	7.9	373	16	Sequence 149274

14	268.5	7.9	802	9	US-09-823-240-2	Sequence 2, Appli
15	267	7.9	280	17	US-10-425-115-231956	Sequence 231956,
16	267	7.9	445	16	US-10-437-963-161789	Sequence 161789,
17	266.5	7.9	469	16	US-10-437-963-136216	Sequence 136216,
18	266.5	7.9	497	16	US-10-437-963-122257	Sequence 122257,
19	266	7.9	501	16	US-10-437-963-161137	Sequence 161137,
20	265	7.8	289	16	US-10-437-963-172302	Sequence 172302,
21	263.5	7.8	563	16	US-10-437-963-198755	Sequence 198755,
22	263	7.8	408	16	US-10-437-963-184023	Sequence 184023,
23	263	7.8	554	16	US-10-437-963-174151	Sequence 174151,
24	262.5	7.8	639	14	US-10-369-493-2275	Sequence 2275, Ap
25	262.5	7.8	1405	16	US-10-437-963-156135	Sequence 156135,
26	261.5	7.7	507	16	US-10-437-963-124879	Sequence 124879,
27	261	7.7	285	16	US-10-437-963-189482	Sequence 189482,
28	261	7.7	1567	9	US-09-835-232-2	Sequence 2, Appli
29	261	7.7	1567	14	US-10-308-485-2	Sequence 2, Appli
30	260.5	7.7	554	16	US-10-437-963-119377	Sequence 119377,
31	258.5	7.6	442	16	US-10-437-963-197867	Sequence 197867,
32	258	7.6	431	16	US-10-437-963-204963	Sequence 204963,
33	256	7.6	500	16	US-10-437-963-139732	Sequence 139732,
34	255.5	7.5	720	14	US-10-342-331-4	Sequence 4, Appli
35	255.5	7.5	1023	10	US-09-893-519A-14	Sequence 14, Appli
36	255	7.5	504	14	US-10-342-331-3	Sequence 3, Appli
37	254.5	7.5	514	16	US-10-437-963-111701	Sequence 111701,
38	254.5	7.5	574	14	US-10-168-097A-76	Sequence 76, Appl
39	254.5	7.5	574	14	US-10-239-431A-38	Sequence 38, Appl
40	254	7.5	533	16	US-10-437-963-167716	Sequence 167716,
41	253.5	7.5	3122	14	US-10-200-562-201	Sequence 201, App
42	253.5	7.5	3122	14	US-10-237-551-201	Sequence 201, App
43	253.5	7.5	3122	14	US-10-237-551-250	Sequence 250, App
44	252.5	7.5	376	16	US-10-437-963-154970	Sequence 154970,
45	252	7.4	903	14	US-10-156-761-11093	Sequence 11093, A

ALIGNMENTS

RESULT 1

US-10-623-429-9
; Sequence 9, Application US/10623429
; Publication No. US20040072152A1
; GENERAL INFORMATION:
; APPLICANT: DAVID M. KOELLE
; APPLICANT: ZHI LIU
; APPLICANT: LAWRENCE COREY
; TITLE OF INVENTION: RAPID, EFFICIENT PURIFICATION OF
; TITLE OF INVENTION: HSV-SPECIFIC T-LYMPHOCYTES AND HSV ANTIGENS IDENTIFIED VIA
; TITLE OF INVENTION: SAME
; FILE REFERENCE: 30967.11USUI
; CURRENT APPLICATION NUMBER: US/10/623,429
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: 60/396,791
; PRIOR FILING DATE: 2002-07-18
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 637
; TYPE: PRT
; ORGANISM: Herpes Simplex Virus 2
US-10-623-429-9

Query Match 100.0%; Score 3385; DB 15; Length 637;
Best Local Similarity 100.0%; Pred. No. 1.1e-199;
Matches 637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MASAEMRERLEAPLPDRAVPYVAGFLALYDSGDPGALDPTVRAALPPENPLPINV 60

Db 1 MASAEMRERLEAPLPDRAVPYVAGFLALYDSGDPGALDPTVRAALPPENPLPINV 60

Qy 61 HRARCEVGRVLA VVNDPRGPFVGLIACVOLERVLETAASAAIFERRGPALSRERLLYL 120

Db 61 HRARCEVGRVLA VVNDPRGPFVGLIACVOLERVLETAASAAIFERRGPALSRERLLYL 120

121 ITNYLPSVSLTKRGDEVPDRTLFAHVALCAIGRRLLGTIVTYDTSDDAAIAPFRHLDP 180
121 ITNYLPSVSLTKRGDEVPDRTLFAHVALCAIGRRLLGTIVTYDTSDDAAIAPFRHLDP 180
181 ATREGVRRRAAEALALAGRTWAPGVEALHTLLSTAVNNMLRDRWSLVAERRRQAGIA 240
181 ATREGVRRRAAEALALAGRTWAPGVEALHTLLSTAVNNMLRDRWSLVAERRRQAGIA 240
241 GHTYLOASEKFKIWGAESAPAPERGKTCAPGAMDTSPAASVPAPQVAVRARQVASSSS 300
241 GHTYLOASEKFKIWGAESAPAPERGKTCAPGAMDTSPAASVPAPQVAVRARQVASSSS 300
301 SSFPAPADMNVPVSAGAPAPPPGAGSGLWIPASHYNQLVTGOSAPRHPLTACGLPAAG 360
301 SSFPAPADMNVPVSAGAPAPPPGAGSGLWIPASHYNQLVTGOSAPRHPLTACGLPAAG 360
361 TVAYGHGPGAGSPHYPPPPPAHPYPCGMLFAGPSPLEAQIAALVGAIAADRQAGGLPAAAGD 420
361 TVAYGHGPGAGSPHYPPPPPAHPYPCGMLFAGPSPLEAQIAALVGAIAADRQAGGLPAAAGD 420
421 HGIRSAKRRRHEVEQPEYDCGRDEPDRFPYYPGEPVDSRRRAARQAASGPHETI 480
421 HGIRSAKRRRHEVEQPEYDCGRDEPDRFPYYPGEPVDSRRRAARQAASGPHETI 480
481 TALVGAVTSLOQELAHMRARTHAPYGPVPGVYHHPHADTETPAQPPRYPAKAVYLP 540
481 TALVGAVTSLOQELAHMRARTHAPYGPVPGVYHHPHADTETPAQPPRYPAKAVYLP 540
541 HIAPGPPPLSGAVPPSYPPVAVTPGAPPLHQPSPAHAPPPPPPGTTPPPPAASLPQPE 600
541 HIAPGPPPLSGAVPPSYPPVAVTPGAPPLHQPSPAHAPPPPPPGTTPPPPAASLPQPE 600
601 APCAAGALVNASSAAHVNVDTARAADLFVSQMMGSR 637
601 APCAAGALVNASSAAHVNVDTARAADLFVSQMMGSR 637

RESULT 2
US-10-214-932-86
; Sequence 86, Application US/10214932
; Publication No. US20030100707A1
; GENERAL INFORMATION:
; APPLICANT: HWANG, Inhwan
; APPLICANT: KIM, Dae Heon
; APPLICANT: LEE, Yong Jik
; TITLE OF INVENTION: SYSTEM FOR DETECTING PROTEASE
; FILE REFERENCE: APB02/US
; CURRENT APPLICATION NUMBER: US/10/214,932
; CURRENT FILING DATE: 2002-08-08
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 86
; TYPE: PRT
; LENGTH: 635
; ORGANISM: Human herpesvirus 1
US-10-214-932-86

Query Match 68.9%; Score 2333; DB 14; Length 635;
Best Local Similarity 74.0%; Pred. No. 4.2e-135;
Matches 477; Conservative 30; Mismatches 120; Indels 18; Gaps 7;

1 MASAEWRLEAPLPRAPYPIVAGFLALYDSGDPGELALDPTVRAALPPENPLPINDV 60
1 MAADAPGRMEEPPLPRAPYPIVAGFLALYDSGDPGELALDPTVRAALPPENPLPINDV 60
61 HRARCEVGRVLA VNDPRGFFVGLIACVOLERVLETAASAIFERRGPGALGREERLLYL 120
61 HRARCEVGRVLA VNDPRGFFVGLIACVOLERVLETAASAIFERRGPGALGREERLLYL 120
121 ITNYLPSVSLTKRGDEVPDRTLFAHVALCAIGRRLLGTIVTYDTSDDAAIAPFRHLDP 180
121 ITNYLPSVSLTKRGDEVPDRTLFAHVALCAIGRRLLGTIVTYDTSDDAAIAPFRHLDP 180

181 ATREGVRRRAAEALALAGRTWAPGVEALHTLLSTAVNNMLRDRWSLVAERRRQAGIA 240
181 ASREGARRLA AEALALSGRTWAPGVEALHTLLSTAVNNMLRDRWSLVAERRRQAGIA 240
241 GHTYLOASEKFKIWGAESAPAPERGKTCAPGAMDTSPAASVPAPQ----VAVRARQVAS 296
241 GHTYLOASEKFKIWGAESAPAPERGKTCAPGAMDTSPAASVPAPQ----VAVRARQVAS 296
297 SSSSSFPAPADMNVPVSAGAPAPPPGAGSGLWIPASHYNQLVTGOSAPRHPLTACGL 356
301 S-----FVLPMPNVPVTSPTAPAPGDPGSGSYLWIPASHYNQLVAGHAAPQPHSARGF 354
357 P-AAGTVAYGHGPGAGSPHYPPPPPAHPYPCGMLFAGPSPLEAQIAALVGAIAADRQAGGLP 415
355 PAAGSVAYGPHGAGLSQHYPPHVAHQTPGVLFSGPSLEAQIAALVGAIAADRQAGGP 414
416 AAAGDHGIRGSAKRRRHEVEQPEYDCGRDEPDRFPYYPGEPVDSRRRAARQAASG 475
415 -AAGDPGVGRSGKRRRYEAGPSESICYDQDEPDADYPYYPGEPVDSRRRAARHSPG 473
476 PHETTALVGAVTSLOQELAHMRARTHAPYGPVPGVYHHPHADTE--TPAQPPRYPAKA 534
474 TNETITALMGAVTSLOQELAHMRARTSAPYGYMTPVAHYRQVGSPEPTTTTHALCPPEA 533
535 VYLPPEPHIAPPGLPSGA--VPPPSYPPVAVTPGAPPLHQPSPAHAPPPPPPGTTPPP 592
534 VYRPPHSAIPGPPGPPASHAPTTPYAAACPPGPPPP---PCPSTQTRAPLPTTEPAFP 590
593 AASLPQPEAPGAEGALVNASSAAHVNVDTARAADLFVSQMMGSR 637
591 AATGSQLPEASNAEAGALVNASSAAHVDVDTARAADLFVSQMMGAR 635

RESULT 3
US-10-779-597-30
; Sequence 30, Application US/10779597
; Publication No. US20040234953A1
; GENERAL INFORMATION:
; APPLICANT: Oregon Health & Science University
; APPLICANT: Wong, Scott W.
; APPLICANT: Axthelm, Michael K.
; APPLICANT: Hansen, Scott G.
; TITLE OF INVENTION: JAPANESE MACAQUE HERPESVIRUS NUCLEIC ACID SEQUENCES AND THEIR USE
; FILE REFERENCE: 178-67426
; CURRENT APPLICATION NUMBER: US/10/779,597
; CURRENT FILING DATE: 2004-02-12
; PRIOR APPLICATION NUMBER: 10/276,524
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: PCT/US01/16274
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/205,652
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30
; LENGTH: 536
; TYPE: PRT
; ORGANISM: Japanese Macaque Herpesvirus
US-10-779-597-30

Query Match 10.4%; Score 351; DB 17; Length 536;
Best Local Similarity 23.3%; Pred. No. 1.8e-13;
Matches 143; Conservative 78; Mismatches 249; Indels 144; Gaps 22;

21 IYVAGFLALYDSGD--PGELALDPTVRAALPPENPLPINDVHRARCEVGRVLA VNDPRG 79
4 VYVGVVDVWSLPKTEKDLYLEPSIVATLLPTNPLPINIEHVPEAHVGHGTLGLFQVTHG 63
80 PFVGLIACVOL----ERVLETAASAIFERRGPGALGREERLLYLITNYPVSLSTKR 135
64 IFCLGKLTSPDFLALASRLAGDSRAAQIQNMP-----RDLPLEMLHTWLPELSLSLHP 119
136 GDEVPPDR--TLFAHVALCAIGRRLLGTIVTYDTSDDAAIAPFRHLDPATREGVRRRAEAE 194

Db 120 EELQDPNHPFAQVSVLCALRRRGSIAVGPDTWVVSKE--DSUTRDEAGKITSKCL 176
Qy 195 LALAGRTPWPCVEALHTLLSTAVNNMLDRWSLVAERRRQAGIAGHTYLOASEKEKIW 254
Db 177 DLCEQVTPPEFAFLETLAKAIDAGFIRDRTDLLTKDGVARVASTYLKASQ----- 231
Qy 255 GAESAPAPERGYKTGAPGAMDTSPAASVPAPQAVRARQVASSSSSSFFAPADMPVSA 314
Db 232 ----SPSSQHG-----GNRDTQMSALPDNDITI-PKSTFLTWVQSSLDHMRNQGRAY 280
Qy 315 SGAPAPPPGDSY-LAIPASHYNQLVTGQAPRHPPLTACGLPAAGTVAYGHGPGAPSP 373
Db 281 VSAP-PSMPATAAYPSMIPP-----PELTV-----P 305
Qy 374 HYPPPAHYP--GMLFAGSPLEAQIAALVGAIAADRQAGGLPAAAGDHGIRGSAKRER 431
Db 306 SYAPPVAPPPFQSAFAPQSPYATYS-----PTYGAQASRH-----QKRK 350
Qy 432 HEVEQPEYDCRDEPDRDPFYYPGGEARPEPRPVDSSRAARQASGPHETITLVGAVTSLQ 491
Db 351 RDVEL-----SDEPVFPG-----EVGIHKDVMAISKNILDIQ 383
Qy 492 QELAHMARHAPYCPYPPVGYHHPHADTETPAOPP--RYPKAVYLPHPHIAPPGL 549
Db 384 ADLRDLKRAASQTSG-----EKDTDRPQPPVQVFSWPQTYASAPYL----- 425
Qy 550 SGAVPPSPYPVAVTPGAPPLHQSP-----AAHAPPPPPPGPTPPPAASLPQPEAP 602
Db 426 --AYQPMY-----PGTDTHLHASQBYQIAQGIQQTQPPPPQASHHAGLATQFVAPAP 477
Qy 603 GAEGALVNASAA 616
Db 478 AAQESAMNAVPSA 491

RESULT 4

US-10-437-963-159995
; Sequence 159995, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT FILING DATE: 2003-05-14
; PRIOR FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 159995
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(534)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_59316C.1.pap
US-10-437-963-159995

Query Match 8.7%; Score 294.5; DB 16; Length 534;
Best Local Similarity 29.3%; Pred. No. 5.2e-10;
Matches 119; Conservative 19; Mismatches 137; Indels 131; Gaps 18;
Qy 256 AESAPAPERGYKTGAPGAMDTSPAASVPAPQAVRARQVASSSS--SSSFAP----- 306

Db 139 AAAAAAPPAR-----APPATSPPPPPPPPPPPPPPPSSSSSSCSATSPPPCPWAAQGT 193
Qy 307 ADMNVSASGAPAPPPPGDGSYLIPASHYNQLVTGQAPRHPPLTAC----- 354
Db 194 AALPPVSSSSXAPPPPPPP-----PXS-----PAPPQPPPPACTRTTSTXPDERR 239
Qy 355 -----GLPAAGTVAYGHGAG-----PSPHY-----PPPAHPYPGMLFAGSPPL 394
Db 240 EABQPRPRPASA-----EGVNHARDVPPPPHRELMPXPPPPPPPPPPPPPPETEOPPP 293
Qy 395 EAQIA-----ALVGAIAADRQAGGLPAAAGDHGIRGSAKRERHEVEQPEYDCG 442
Db 294 SAYLSPSHSGYSSAPPLPQATAAAVRAPFRPAPALPH--CGCA--RQHERLRP----- 344
Qy 443 RDEPDRDPFYYPGEARPEPRPVDSSRAARQASGPHETITLVGAVTSLQELAHMARHARTH 502
Db 345 -----PRXPPPPSPSAVPAASP-----SVLPPPAS 372
Qy 503 APYGP-----YPPVGPYHHPHADTETPAOPPRYPYPAKAVYLPHPHIAPPGLSGAVPPPSY 558
Db 373 PPXPVPPVLPPLPPPCPSPLPAPLRSPSPPTXPAALHTLP-----PPTPCAGKPPPPX 427
Qy 559 PPVAVTPGAPPLHQ--PSPAHAHPPPPPPGPTPPPAASLPQPEAP 602
Db 428 PPPPPPPXPRPADSUPPPSPSAPPSPPP-PPPPPPPLPPPPRAP 472

RESULT 5

US-10-369-493-3962
; Sequence 3962, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 3962
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Neurospora crassa
US-10-369-493-3962

Query Match 8.6%; Score 292; DB 14; Length 639;
Best Local Similarity 24.0%; Pred. No. 9e-10;
Matches 126; Conservative 28; Mismatches 190; Indels 180; Gaps 20;
Qy 176 RLHLP-ATREGVRRAAEAELALAGRT-----NAPGVEALT-----HTLLST 216
Db 177 KHASPTATPPSPRAAAPACRMRAGSTNGINGYKYPSEFATLDAPDPLWREHFDDDLKQ 236
Qy 217 AVNNMLRDRWSLVAE--RRRQAGIAGHTYLOASEKFIWGAESAPAPERGYKTGAPCAM 274
Db 237 GLDDDFIRENQDFIIEFLKQEQAKQGIHT-----STPPP-----PPPPPSAP 278
Qy 275 DTSFAASVPAPQAVRARQVASSSSSSFFAPADMPVSAAGAPAPPPGDSGLWLPAS 334
Db 279 PNGPAMRAPPPPPPPAAAPRSVSES-----ITPSTSRRGVPPPPPP-----PAR 321
Qy 335 HYNQVLT--GOSAPRHPPLTACGLPAAGTVAYGH--PGAGSPHYPPPPRAHP----- 382
Db 322 RSGKLDTENHQEPAPPPRFAPVPIADAGKFAHSDPPRHTPSAPGPPPPPPKTPLEDQ 381
Qy 383 YPGMLFAGSPFLBQAIAALVGAIAADRQAGGLPAAAGDHGIRGSAKRERHEVEQPEYDCG 442

Db 382 DPSQRESVPPFTGQ----- 396
QY 443 RDEPDRDFYYGCEARPERPRVDSSRAARQAGSPHETITIALVCAVTSLOQELAHMARH 502
Db 397 -----RSVFPFPPSRSSVPPPPRNSAAQPPLP-----PKAP 429
QY 503 APYGVPPVGVYHHHADTETAQPPRYPAKAVYLPPLP-----HIAPGPPPLSGAVPPPS 557
Db 430 GRAPPLFPASSRPPMLPRTSRFAPPOAPPLPTSNAPPPPLPATAQAPPPPLPATSAPPP 489
QY 558 YPFVAVTPGAPPL-----HQFSP-----AHAPPPPPPPG-----PTPPPA 593
Db 490 PPPAPPAP-PAPPLPAHAHAPPPPPMPMPAPSGGAPPPPPPPPGMGVPPPPPPPP 548
QY 594 ASLPQPEAP-----GAEGALVNASSAAHV-----NVDTARAAD 627
Db 549 GGMPPPPAPAPFPVDGSRSAVLGDKTAGGIRALKKVDRSQIRD 592

RESULT 6

US-10-408-765A-2285
; Sequence 2285, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale B.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2285
; LENGTH: 825
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2285

Query Match 8.4%; Score 284; DB 16; Length 825;
Best Local Similarity 24.7%; Pred. No. 3.7e-09;
Matches 159; Conservative 58; Mismatches 230; Indels 196; Gaps 30;

QY 35 PGELALDPDTVRAALPPENPLPINVDHRARCEVGRVIAV-----VNDPRGPF 81
Db 185 PG-LGLDPNK-----PPPLFSTVNAEPLGLIOGPHQAAPPPPPPPPPAPASEPKGL 238
QY 82 FVGLIACVOLERVLETAASAAIFERRGPALSREERLLYLITNYPGSVLSLTKRRGDVEVP 141
Db 239 TSPIFCTFKPKLLKT--SFHLLRRDDPFQPKL-----YAOEYFEADEDKADVPA 291
QY 142 DRTLFAHVALCAIGRLGTIVTYDTSLSDAIAAFRHLD-----PAT 182
Db 292 DTRLNP-----RRLPLVSSCRS-REALSPLGDI DFCPPNPGPDGPRRRGRKPTAK 342
QY 183 REGVREAAEAL-ALAGTWTAGVEALHTLLSTAVNNMLRDRWSLVAERRRQAG 241
Db 343 RDGPPRGRPRIRPLEVTTAGPASASTPT---DGAKKPRGR-----GGRKABEAG 394
QY 242 HTYLOASEFKTWGAESAPAPERGYKTGAPAMD-----TSPAASVPAQVAVRAAROVASS 297
Db 395 GTLEPLKPLKI--KLSVPKAGEGLTSSGDAISGTDHNSLDSLTREKIEAKIKEY--- 449
QY 298 SSSSSPPADNMNPSASGAPAPPPPGDGSYLWIPASHYNQLVTGOSAPRHPPLTACGLP 357
Db 450 -----EEKQPMKSG-----FMAASFLDFLKSCK-----RHPPLYQAGLT 483
QY 358 AAGTVAIGHGAGPSHYPP-----PPAHYPYGMPLFAGPSPLEAQIAALVCAI-A 406

Db 484 -----PPLSPKSVPPSPVARGLQPOQPATP-----AVPFPSPSGAGLGGALEA 528
QY 407 ADROQAGGLPAAAGDHGIRGSAKRRRHEVEQBYDCGRDEPDRDFPYYPGEARPEPRPVD 466
Db 529 ABSEGLGL-----GCPSPCKRLDELK-----RNLETLPSFSDEEDSVAK 569
QY 467 RRAARQASGPHETITIALVCAVTSLOQELAHMARHAPYGPYPV-GPVYHHHADTETA 525
Db 570 NRDLQ-----ESTSSAISALDD-----PPLAGP-----KDTSTPD 599
QY 526 QPPRYPAKAVYLPPL-----HIAPGPPPLSGAVPPPSYPVAVTPGAPPLHQPSPA 577
Db 600 GPFLAPAAAVGFPPLPGLPSANSNGTPEPPLLEKPPPT-PPAPTPOQPQPPPPPPQ 658
QY 578 HAHF-----PPPPPGTTPPPAASLPQPEAPGAEGA 608
Db 659 PALPSPPLVAPTPSSPPPLPLPPPPPA-MSPSPPPPPPPAAA 700

RESULT 7

US-10-437-963-191498
; Sequence 191498, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 191498
; LENGTH: 483
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(483)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MBT4530_87810C.1.pap
US-10-437-963-191498

Query Match 8.3%; Score 281.5; DB 16; Length 483;
Best Local Similarity 28.1%; Pred. No. 2.9e-09;
Matches 99; Conservative 12; Mismatches 130; Indels 111; Gaps 15;

QY 303 FPAPADMNPSASGAPAPPPPGDGSYLWIPASHYNQLVTGOSAPRHPPTACGLPAAGTV 362
Db 85 YPPFX---PPYPGXYPYPPPPYPYPPYPPXK---RPTSDPRPRQPP--PCPPXPPPP 136
QY 363 AVGHGAGPSH-----YPPPPAHYPYGMPLFAGPSPLEAQIAALV 402
Db 137 PYPPPPPYPPYPPPPYPPWPGXPYPPYPLPYPPXSPYFGWYPPYP----- 187
QY 403 GAIAADROAGGLPAAAGDHGIRGSAKRRRHEVEQYDCGRDEPDRDFYYP----- 454
Db 188 -----GWPE-----YPPYPPYAGRYPPYPPYPPYPPXDXIH 221
QY 455 -----GEARPEPRVDSRRARQASGPHETITIALVCAVTSLOQELAHMARHAPYGPYPV 511
Db 222 LFTSDRPPPPPPRRPXAAPPAPPPGPP-----PXTSPDPLVGRASPALP--PPPPPP 273
QY 512 GPY-----HHPHADTETAQPPRYPAKAVYLPPLPPPHIAPPGGPLSGAVP- 554


```

RESULT 9
US-10-437-963-142229
; Sequence 142229, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 142229
; LENGTH: 426
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(426)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_43255C.1.pap
US-10-437-963-142229

Query Match      8.1%; Score 275.5; DB 16; Length 426;
Best Local Similarity 27.9%; Pred. No. 5.9e-09;
Matches 107; Conservative 26; Mismatches 185; Indels 65; Gaps 12;

QY  258  SAPAPERGYKTGACGAMDTSPAAS-----VPAPQVAVRARQVASSSSSSFFAPADNPFV 312
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
      60  AATGFGRGPDPAAPTVAADPTTTTSLPLPLPPSPPLDLAAPAADLASSPPPPPHAVRT 119

QY  313  SASGAPAPPBGDCGSLWIPASHYNQLVTGQSAPRRPPLTACGLPAAGTVAVGH-----PG 368
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
      120  SPSSXPXPPPP-----SPVIPPQRPPXPASPPRPAAGPRQPRPTVVGCTSMRHGCLSP 176

QY  369  AGSPSPHYPPPAHPYPCGMLFAGSPSPLEAQIAALVGAIAADROAGGLPAAAGDHGIRGSAK 428
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
      177  AEPSCXPPPPAVAPQPQTRATPAPQAPLXA-----AAGSSRGDPAAHTAAVAAPTAT 231

QY  429  RRRHEVQPEYDCGRDEPRDF-----PYYPGEAR-PEPRPVDSRRRAARQASGPHETI 480
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
      232  PPLPPLPPPPSPPLDLAAPTTLASLSQPPPPPTKKKTPSPXFPVPPRHVTPPPPRPH--- 288

QY  481  TALVGATYSLQOELLAHMRARTHAPYGPYPVPGYHHPHADTETPAQPPRYPAKAVLPP- 539
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
      289  -----VPSIPP-----PPXPPPPPLPLFWPPPPPLLSPAWXPSPSPPTPPARLPPX 332

QY  540  PHIAPPGPPLISGAV--PPSPYPPVAVTPGAPPLHQPSPAHAHPPP--PPPGGTP----- 590
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
      333  PSTSRPPPPRTPLPLDPPSPPPPLFLPPPPPPXSPVNPFRHSQPTPHSPSPSPXSLPLP 392

QY  591  -----PPAASLFPQPEAP 602
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
      393  LVAPPHRRRRSSPPAVRRRRSPAP 415

RESULT 10
US-10-437-963-126337
; Sequence 126337, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.

```

; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 126337
; LENGTH: 506
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(506)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_28894C.1.pep
US-10-437-963-126337

Query Match 8.1%; Score 273; DB 16; Length 506;
Best Local Similarity 29.2%; Pred. No. 1e-08;
Matches 109; Conservative 18; Mismatches 130; Indels 116; Gaps 19;
QY 270 AFGAMDTSPAASVPAPQAVRA---RQVASS-----SSSSGFAPADMN-----P 311
DB 30 SFIAVIASSPVRAPLARVPACRRRVASSHPAALCVLSPSPVPRXPPSSPPPPAP 89
QY 312 VSASGAPAPPPGDSYLIWIPASHYNQVLTGOSAPRHPLTACGLPAACTVAYHGPAGP 371
DB 90 XPRACAPPPPPPD-----VV-----PPPSAPXPPAGPPARRPP---P 125
QY 372 SPHP-----PPPA---HPYRCMLFAGSPLEAQIAALVGAIAADRAAGGLPAAAG 419
DB 126 SPLRDLGREGGEPFPAASRRPAPP---ASHPPPAR-----SGG 163
QY 420 DHGIRGSAKRHEVEQPYDCGRDEPDRDFVYVYGEARPEP-RP-VDSRRAARQASGP- 476
DB 164 XEGEKPPATPRTXKRPSPDAVRSPDGIRSPSPSAGLPRLRDLGREGGEPXPXP 223
QY 477 --HETITLVAVTSLQELAHMRARTHAPYGVPPVGVYHHP-----HADTETAPQPR 529
DB 224 CARPPVRALDAVR-----PSPPASPLPPPARSGGEGXEPXPXPPL 265
QY 530 YPAKAVYLPPIHAPGPPPLSGAVRPPSPVPPVAVTGPAPPLHQSPAHAPPPPPPGPT 589
DB 266 PFPQ-----PPRLAPQRPD--PAVPPP--PPPREGPPPPPPPPPPPP-----PPPPPPPPPL 313
QY 590 PPPAASLQPEAP 602
DB 314 PPPLAPRCEPPLP 326

RESULT 11
US-10-437-963-179715
; Sequence 179715, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 179715
; LENGTH: 521
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(521)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_7714C.1.pep
US-10-437-963-179715

Query Match 8.0%; Score 271.5; DB 16; Length 521;
Best Local Similarity 27.1%; Pred. No. 1.3e-08;
Matches 95; Conservative 20; Mismatches 117; Indels 119; Gaps 17;
QY 256 AESAPAPERGYKTGAPGAMDTSPAASVPAPQAVRQVASSSSSSSFFA-PADMNVPVA 314
DB 20 AATSPSP-----PPGRAGNVPTSAREPSVT-----PTWPTTGPPTP 57
QY 315 SGAPAPPPPGDSY--LWIPASHYNQVLTGOSAPRH---PPLTACGLPAACTVAYHGPAG 369
DB 58 SARPPPLPPTTSAHPEAATPPPPHAPPAPPOQXPSPHSPPLPTLTVPP-----HRRH 110
QY 370 GPSPHYPPPAHPYPCMLFAGSPLEAQIAALVGAIAADRAAGGLPAAAGDHGIRGSAKR 429
DB 111 YPNHLPFPQPPLB---YTPP-----RRSPPR 135
QY 430 RHHEVEQPYDCGRDEPDRDFVYVYGEARPEPRVDSRRAARQASGPHETITLVAVTIS 489
DB 136 SRPPTTP-----SRPPPPPAHPPPPPQPTP---RR--HHPTTTPH-----APAAPT 179
QY 490 LQELAHMRARTHAPYGVPPVGVYHHPHADTETPAQPRYPKAVYLPPIHAPPGPPL 549
DB 180 XPQ-----RTTP--PPPP-----PPPDTSRTPPTPPRAPPTPPRPGPPPPPL 224
QY 550 SGAVPPSPYPVAVTGPAPPLHQSPAHAPPPPPPGTTPPPAASLQPE 600
DB 225 ----PPPPPPAPPPEG-----PPPPPPXPPQPTPSLPPKE 257

RESULT 12
US-10-437-963-195106
; Sequence 195106, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 195106
; LENGTH: 606
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(606)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_91087C.1.pep

Search completed: January 10, 2005, 19:31:11
Job time : 150 secs

```
QY 423 IRGSAKRREHVEQPYDCGRDEPDYFYFGEARPEP-RVVDSSRAARQASGPHETIT 481
Db 471 -----CG-----SQASPPPGTPLASTFSSKPSVLSPSPSAG 500
QY 482 ALVGAVTSLQOELAH-----MRARTHAPYGPYP---VGPYHHPHADTETPAOPPRYP 531
Db 501 APASAEPLNPELGDSSASEPGLQASQPAESFTTQGLVLP-----PAPPP--- 547
QY 532 AKAVYLPPPHIAPPGPPLSGAVPPSPYPPVAVTPGPAFLHQFSPAHAHPPPPPPGCP--- 588
Db 548 -----PPP--LPSCGAYASALPPP-----PGPPPP--PPLFSTGPPPPPPPPPPPLP 589
QY 589 --TPPPAASLPOPEAFGEAGALVNASSA 615
Db 590 NOAPPPP-----PPPPAPPLPASGIFSGSTS 615

RESULT 15
US-10-425-115-231956
; Sequence 231956, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 231956
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(280)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_143137C.1.pap
US-10-425-115-231956

Query Match 7.9%; Score 267; DB 17; Length 280;
Best Local Similarity 27.8%; Pred. No. 1.3e-08;
Matches 84; Conservative 6; Mismatches 84; Indels 128; Gaps 12;

QY 304 PAPADMNVPYSACAPAPPPGDSYLIWIPASHYNQLVTCQSAPRHPPLTACGLPAAGTVA 363
Db 65 PPPFHQHRTSPGPPPPPPPPG-----PPPHPPATPPPPXPPP----- 100
QY 364 YGHPGAGSPGPHYPP-PPAHPYPGMLFAGPSPLEAQIALVGAIAADRAQAGLPAAGDHG 422
Db 101 ----PPPPXPHPPPPQPPNP-PPTHPPPTP----- 127
QY 423 IRGSAKRREHVEQPYDCGRDEPDYFYFGEARPEP-RVVDSSRAARQASGPHETITA 482
Db 128 -----PPRPPPPPPXPTSPPPPPAPSR-----GP----- 152
QY 483 LVGAVTSLQOELAHMRARTHAPYGPYPVGPYHHPHADTETPAQPPRYPAKAVYLLPPPHI 542
Db 153 -----PPPPXPPNPPPPHP-----TPPPPPPPPPPPPPPPPPPS 186
QY 543 APPGPPLSGAVPPSPYPPVAVTPGPAFLHQFSPAHAHF--PPPPPGTPPPAASLPQPE 600
Db 187 SPPPPPL-----PPP--PPLLPSPSPHPPPTPPPPH-HRQPPPPPPPPPPNPTPPPT 239
QY 601 AP 602
Db 240 PP 241
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2005, 18:56:04 ; Search time 160 Seconds
(without alignments)
1428.191 Million cell updates/sec

Title: us-10-623-429-9

Perfect score: 3385

Sequence: 1 MSAENRERLEAPLPDRAVP.....VNVDTARAADLFVQMWGSR 637

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq_23Sep04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3385	100.0	637	6	Aae36428 HSV-2 UL2
2	3385	100.0	637	8	Adj77800 Herpes si
3	3370.5	99.6	638	2	Aar71017 HSV-2 pro
4	3359.5	99.2	642	2	Aaw72124 HSV-2 str
5	3351.5	99.0	636	2	Aar64767 HSV-2 pro
6	3267.5	96.5	657	2	Aaw72192 HSV-2 str
7	2333	68.9	635	2	Aar71031 HSV-1 UL2
8	2333	68.9	635	7	Adl18166 Human her
9	2329	68.8	635	2	Aar28634 UL26 prot
10	2326	68.7	635	2	Aar28651 UL26 prot
11	2323.5	68.6	636	2	Aar28636 UL26 prot
12	2322	68.6	635	2	Aar28647 UL26 prot
13	2321	68.6	635	2	Aar28652 UL26 prot
14	2319	68.5	626	2	Aar28640 UL26 prot
15	2319	68.5	635	2	Aar28650 UL26 prot
16	2318.5	68.5	636	2	Aar28637 UL26 prot
17	2318.5	68.5	636	2	Aar28635 UL26 prot
18	2318	68.5	635	2	Aar28649 UL26 prot
19	2306	68.1	635	2	Aar28648 UL26 prot
20	2241	66.2	615	2	Aar28642 UL26 prot
21	2209	65.3	603	2	Aar28641 UL26 prot
22	2017	59.6	514	2	Aar28643 UL26 prot
23	2014.5	59.5	448	2	Aaw72048 HSV-2 str
24	1826	53.9	331	2	Aar64769 HSV-2 ISP
25	1820	53.8	329	6	Aae36429 HSV-2 UL2

RESULT 1
AAE36428

ID AAE36428 standard; protein; 637 AA.

XX AC AAE36428;

DT 07-AUG-2003 (first entry)

XX DE HSV-2 UL26 full-length antigen.

XX KW Herpes simplex virus; HSV; infection; vaccine; therapy; UL26; antigen.
XX OS Herpes simplex virus type 2.
XX PN WO2003020108-A2.
XX PD 13-MAR-2003.
XX PF 27-AUG-2002; 2002WO-US027341.
XX PR 04-SEP-2001; 2001US-0317159P.
PR 17-APR-2002; 2002US-0373429P.
XX (CORI-) CORIXA CORP.
PA Hosken NA, Day CH;
PI WPI; 2003-290135/28.
DR N-PSDB; AAD55178.
XX Novel isolated polypeptide comprising immunogenic portion of a herpes simplex virus antigen, useful for detecting herpes simplex virus infection in a subject, and for treating the virus infection in a patient.
XX Claim 1; Page 110-112; 114pp; English.
XX The invention relates to polypeptides comprising an immunogenic portion of herpes simplex virus (HSV) antigen and to nucleic acid molecules encoding polypeptides. Polypeptides of the invention are useful for detecting and treating HSV infection in a patient. Polynucleotides of the invention are useful as diagnostic reagents for detecting HSV infection in a patient and also as probes or primers. The invention is used to prepare vaccines. The present sequence is HSV-2 strain HG52 UL26 full-length antigen
XX Sequence 637 AA;

ALIGNMENTS

26	1432	42.3	350	2	AAW10155	Aaw10155 HSV-1 (F)
27	1432	42.3	350	2	AAW09967	Aaw09967 HSV-1 (F)
28	1429	42.2	350	2	AAW09960	Aaw09960 HSV-1 (F)
29	1429	42.2	350	2	AAW09958	Aaw09958 HSV-1 (F)
30	1429	42.2	350	2	AAW09971	Aaw09971 HSV-1 (F)
31	1429	42.2	350	2	AAW09973	Aaw09973 HSV-1 (F)
32	1426	42.1	350	2	AAW09956	Aaw09956 HSV-1 (F)
33	1426	42.1	350	2	AAW09962	Aaw09962 HSV-1 (F)
34	1426	42.1	350	2	AAW09975	Aaw09975 HSV-1 (F)
35	1426	42.1	350	2	AAW09969	Aaw09969 HSV-1 (F)
36	1424	42.1	350	2	AAW09963	Aaw09963 HSV-1 (F)
37	1424	42.1	350	2	AAW09985	Aaw09985 HSV-1 (F)
38	1424	42.1	350	2	AAW09976	Aaw09976 HSV-1 (F)
39	1424	42.1	350	2	AAW09978	Aaw09978 HSV-1 (F)
40	1423	42.0	350	2	AAW09959	Aaw09959 HSV-1 (F)
41	1423	42.0	350	2	AAW09964	Aaw09964 HSV-1 (F)
42	1423	42.0	350	2	AAW09977	Aaw09977 HSV-1 (F)
43	1423	42.0	350	2	AAW09972	Aaw09972 HSV-1 (F)
44	1422	42.0	350	2	AAW09957	Aaw09957 HSV-1 (F)
45	1422	42.0	350	2	AAW09961	Aaw09961 HSV-1 (F)

Query Match		100.0%;	Score 3385;	DB 6;	Length 637;	
Best Local Similarity		100.0%;	Pred. No. 4e-209;			
Matches 637;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	1	MASAEERRELEAPLPDRAVPIYVAGFLALYDSGDPGELALDDPTVRAALPPENPLPINVD	60			
DB	1	MASAEERRELEAPLPDRAVPIYVAGFLALYDSGDPGELALDDPTVRAALPPENPLPINVD	60			
QY	61	HRARCEVGRVLAVNDPRGPFVGLIACVQLSERVLETAASAAIFERRGPALSRERLLYL	120			
DB	61	HRARCEVGRVLAVNDPRGPFVGLIACVQLSERVLETAASAAIFERRGPALSRERLLYL	120			
QY	121	ITNYLPSVLSLTKRGDEVPDRTLFAHVALCAIGRRLGITVYDTSDDAATAPRHLDP	180			
DB	121	ITNYLPSVLSLTKRGDEVPDRTLFAHVALCAIGRRLGITVYDTSDDAATAPRHLDP	180			
QY	181	ATREGVRRERAEALAGRTWAPGVEALHTLLSTAVNNMLRDRWSLVAERRRQAGIA	240			
DB	181	ATREGVRRERAEALAGRTWAPGVEALHTLLSTAVNNMLRDRWSLVAERRRQAGIA	240			
QY	241	GHTYLOASEKFKIWGAESAPAPERGYKTGAPGAMDTSPAASVPAPQVAVRARQVASSSSS	300			
DB	241	GHTYLOASEKFKIWGAESAPAPERGYKTGAPGAMDTSPAASVPAPQVAVRARQVASSSSS	300			
QY	301	SSFPAPADMNPVSASGAPAPPPPGDGSYLWIIPASHYNQLVTGOSAPRHPPLTACGLPAAG	360			
DB	301	SSFPAPADMNPVSASGAPAPPPPGDGSYLWIIPASHYNQLVTGOSAPRHPPLTACGLPAAG	360			
QY	361	TVAYGHGAGSPHYPPPPAHYPGMLFAGSPLEAQIAALVGAIAADRQAGGLPAAAGD	420			
DB	361	TVAYGHGAGSPHYPPPPAHYPGMLFAGSPLEAQIAALVGAIAADRQAGGLPAAAGD	420			
QY	421	HGIRSAKRRRHEVQPEYDCGRDEPDRDFPYYPGEARPEPPVDSRRRAARQASGPHETI	480			
DB	421	HGIRSAKRRRHEVQPEYDCGRDEPDRDFPYYPGEARPEPPVDSRRRAARQASGPHETI	480			
QY	481	TALVGAVTSLOQELAHMRARTHAPYGPYPVGYHHPHADTTTPAQPPRYPAKAVYLP	540			
DB	481	TALVGAVTSLOQELAHMRARTHAPYGPYPVGYHHPHADTTTPAQPPRYPAKAVYLP	540			
QY	541	HIAPGPPPLSGAVPPSPYPVAVTPGPAPLHQPSPAHAHPPPPPPGPTPPPAASLPQPE	600			
DB	541	HIAPGPPPLSGAVPPSPYPVAVTPGPAPLHQPSPAHAHPPPPPPGPTPPPAASLPQPE	600			
QY	601	APGAEAGALVNASSAAHVNVDTARAADLFVSOQMGSR	637			
DB	601	APGAEAGALVNASSAAHVNVDTARAADLFVSOQMGSR	637			
RESULT 2						
ADJ77800						
ID	ADJ77800 standard; protein; 637 AA.					
XX	AC					
XX	ADJ77800;					
DT	06-MAY-2004 (first entry)					
XX	Herpes simplex virus 2 UL7 protein.					
DE	viricide; gene therapy; antigen; skin; lymphocyte;					
KW	peripheral blood mononuclear cell;					
KW	cutaneous lymphocyte-associated antigen; HSV infection.					
XX	Human herpesvirus 2.					
OS	WO2004009021-A2.					
PN	29-JAN-2004.					
XX	18-JUL-2003; 2003WO-US022560.					
PF	18-JUL-2002; 2002US-0396791P.					
PR						
XX						

PA	(UNIW) UNIV WASHINGTON.					
XX	Koelle DM, Liu Z, Corey L;					
XX	WPI; 2004-123280/12.					
DR	Identifying an active antigen of a virus that attacks skin for e.g.					
XX	treating herpes simplex virus, comprises isolating lymphocytes from					
PT	peripheral blood mononuclear cells that express cutaneous lymphocyte-					
PT	associated antigen.					
XX	Claim 9; SEQ ID NO 9; 123pp; English.					
XX	The invention relates to a method of identifying an immunologically					
CC	active antigen of a virus that attacks skin comprises: (1) isolating					
CC	lymphocytes from peripheral blood mononuclear cells (PBMC) that express					
CC	cutaneous lymphocyte-associated antigen (CLA); (2) identifying a CLA-					
CC	positive lymphocyte that selectively kills cells infected with the virus;					
CC	and (3) determining the identity of the antigen present in the lymphocyte					
CC	identified in (2). The method is useful for identifying an					
CC	immunologically active antigen of a virus that attacks skin for preparing					
CC	a composition for treating or preventing an HSV infection. This sequence					
CC	corresponds to the UL7 protein from Herpes simplex virus type 2.					
XX						
SQ	Sequence 637 AA;					
Query Match						
Best Local Similarity 100.0%; Score 3385; DB 8; Length 637;						
Matches 637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;						
QY	1	MASAEERRELEAPLPDRAVPIYVAGFLALYDSGDPGELALDDPTVRAALPPENPLPINVD	60			
DB	1	MASAEERRELEAPLPDRAVPIYVAGFLALYDSGDPGELALDDPTVRAALPPENPLPINVD	60			
QY	61	HRARCEVGRVLAVNDPRGPFVGLIACVQLSERVLETAASAAIFERRGPALSRERLLYL	120			
DB	61	HRARCEVGRVLAVNDPRGPFVGLIACVQLSERVLETAASAAIFERRGPALSRERLLYL	120			
QY	121	ITNYLPSVLSLTKRGDEVPDRTLFAHVALCAIGRRLGITVYDTSDDAATAPRHLDP	180			
DB	121	ITNYLPSVLSLTKRGDEVPDRTLFAHVALCAIGRRLGITVYDTSDDAATAPRHLDP	180			
QY	181	ATREGVRRERAEALAGRTWAPGVEALHTLLSTAVNNMLRDRWSLVAERRRQAGIA	240			
DB	181	ATREGVRRERAEALAGRTWAPGVEALHTLLSTAVNNMLRDRWSLVAERRRQAGIA	240			
QY	241	GHTYLOASEKFKIWGAESAPAPERGYKTGAPGAMDTSPAASVPAPQVAVRARQVASSSSS	300			
DB	241	GHTYLOASEKFKIWGAESAPAPERGYKTGAPGAMDTSPAASVPAPQVAVRARQVASSSSS	300			
QY	301	SSFPAPADMNPVSASGAPAPPPPGDGSYLWIIPASHYNQLVTGOSAPRHPPLTACGLPAAG	360			
DB	301	SSFPAPADMNPVSASGAPAPPPPGDGSYLWIIPASHYNQLVTGOSAPRHPPLTACGLPAAG	360			
QY	361	TVAYGHGAGSPHYPPPPAHYPGMLFAGSPLEAQIAALVGAIAADRQAGGLPAAAGD	420			
DB	361	TVAYGHGAGSPHYPPPPAHYPGMLFAGSPLEAQIAALVGAIAADRQAGGLPAAAGD	420			
QY	421	HGIRSAKRRRHEVQPEYDCGRDEPDRDFPYYPGEARPEPPVDSRRRAARQASGPHETI	480			
DB	421	HGIRSAKRRRHEVQPEYDCGRDEPDRDFPYYPGEARPEPPVDSRRRAARQASGPHETI	480			
QY	481	TALVGAVTSLOQELAHMRARTHAPYGPYPVGYHHPHADTTTPAQPPRYPAKAVYLP	540			
DB	481	TALVGAVTSLOQELAHMRARTHAPYGPYPVGYHHPHADTTTPAQPPRYPAKAVYLP	540			
QY	541	HIAPGPPPLSGAVPPSPYPVAVTPGPAPLHQPSPAHAHPPPPPPGPTPPPAASLPQPE	600			
DB	541	HIAPGPPPLSGAVPPSPYPVAVTPGPAPLHQPSPAHAHPPPPPPGPTPPPAASLPQPE	600			
QY	601	APGAEAGALVNASSAAHVNVDTARAADLFVSOQMGSR	637			
DB	601	APGAEAGALVNASSAAHVNVDTARAADLFVSOQMGSR	637			

RESULT 3
 AAR71017
 ID AAR71017 standard; protein; 638 AA.
 XX
 AC AAR71017;
 XX
 DT 25-MAR-2003 (revised)
 DT 02-OCT-1995 (first entry)
 XX
 DE HSV-2 protease.
 XX
 KW Herpes simplex virus; HSV-2; capsid; UL26 gene.
 XX
 OS Herpes simplex virus type 2.
 XX
 PN W09506055-A1.
 XX
 PD 02-MAR-1995.
 XX
 PF 19-AUG-1994; 94WO-US009303.
 XX
 PR 20-AUG-1993; 93US-00110522.
 PR 23-JUN-1994; 94US-00264537.
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 XX
 PI Dilella AG, Debouck CM;
 XX
 DR WPI; 1995-106803/14.
 DR N-PSDB; AAQ84671.
 XX
 PT New herpes simplex virus (HSV)-2 protease and capsid protein - used to develop prods. for use in the diagnosis and treatment of HSV-2 infections.
 XX
 PS Claim 1; Fig 1; 51pp; English.
 XX
 -CC The sequence is the product of the herpes simplex virus type 2 gene UL26, which is the HSV-2 protease. The protein can be used in the diagnosis and treatment of HSV-2 infections. See also AAR71018-31. (Updated on 25-MAR-2003 to correct PN field.)
 CC
 CC 2003 to correct PN field.)
 XX
 SQ Sequence 638 AA;

Query Match 99.6%; Score 3370.5; DB 2; Length 638;
 Best Local Similarity 99.7%; Pred. No. 3.4e-208;
 Matches 636; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 MASAEMRERLEAPLDRAVPIYVAGFLALYDSDGDELALDPTVRAALPPENPLINVD 60
 DB 1 MASAEMRERLEAPLDRAVPIYVAGFLALYDSDGDELALDPTVRAALPPENPLINVD 60
 QY 61 HRARCEVGRVLAIVNDPRGPFVGLIACVQLERVLETAASAAIFERRGPALESREERLLYL 120
 DB 61 HRARCEVGRVLAIVNDPRGPFVGLIACVQLERVLETAASAAIFERRGPALESREERLLYL 120
 QY 121 ITNLPVSLSTKRGDEVPDRTLFAHVALCAIGRRGLTIVTYDTSLOAAIAPFRHLDP 180
 DB 121 ITNLPVSLSTKRGDEVPDRTLFAHVALCAIGRRGLTIVTYDTSLOAAIAPFRHLDP 180
 QY 181 ATREGVRRRAAEALAGTAPGVEALHTLLSTAVNNMLRDRWSLVAERROAGTA 240
 DB 181 ATREGVRRRAAEALAGTAPGVEALHTLLSTAVNNMLRDRWSLVAERROAGTA 240
 QY 241 GHTYLOASEKFKIWGAEAPAPRGYKTAGPGAMDTSPAASVPAPQVAVRQVA-SSSS 299
 DB 241 GHTYLOASEKFKIWGAEAPAPRGYKTAGPGAMDTSPAASVPAPQVAVRQVASSSS 300
 QY 300 SSSFPAPADNPNVSAGAPAPPPGSGSYLWIPASHYNQLVTQSQAPRHPPLTACGLPAA 359
 DB 301 SSSFPAPADNPNVSAGAPAPPPGSGSYLWIPASHYNQLVTQSQAPRHPPLTACGLPAA 360

QY 360 GTVAYGHGAGSPSPHYPPPAHPYPGMLFAGSPLEAIAALVGAIAADROAGGLPAAAG 419
 DB 361 GTVAYGHGAGSPSPHYPPPAHPYPGMLFAGSPLEAIAALVGAIAADROAGGLPAAAG 420
 QY 420 DHGIRGSAKRRRHEVEQPEYDCGRDEPDPRFPYYPGSEARPEPRPVDSSRAARQASGPHET 479
 DB 421 DHGIRGSAKRRRHEVEQPEYDCGRDEPDPRFPYYPGSEARPEPRPVDSSRAARQASGPHET 480
 QY 480 ITALVGAVTSLQOELAHMRARTHAPYGPYPVPGYHHPHADTETPAOPPRYPKAVVLP 539
 DB 481 ITALVGAVTSLQOELAHMRARTHAPYGPYPVPGYHHPHADTETPAOPPRYPKAVVLP 540
 QY 540 PHIAPPGPLSGAVPPSPYPVAVTPGPAPPLHQPSPAHAPHPPPPGTPPPPAASLPQP 599
 DB 541 PHIAPPGPLSGAVPPSPYPVAVTPGPAPPLHQPSPAHAPHPPPPGTPPPPAASLPQP 600
 QY 600 EAPGAEAGALVNASSAAHVNVDTARAADLFVSQWMSR 637
 DB 601 EAPGAEAGALVNASSAAHVNVDTARAADLFVSQWMSR 638

RESULT 4

AAW72124
 ID AAW72124 standard; protein; 642 AA.
 XX

AC AAW72124;
 XX

DT 18-DEC-1998 (first entry)
 XX

DE HSV-2 strain SB5 Contig ID 15 ORF#11 protein.
 XX

KW HSV-2 strain SB5; immunological response induction; therapy;
 XX antiviral identification; viral protein inhibitor.

OS Herpes simplex virus 2.
 XX

PN W09820016-A1.
 XX

PD 14-MAY-1998.
 XX

PF 31-OCT-1997; 97WO-US020016.
 XX

PR 04-NOV-1996; 96US-0030279P.
 XX

PR 09-JUN-1997; 97US-0049018P.
 XX

PA (SMIK) SMITHKLINE BEECHAM CORP.
 XX

PI Esser KM, Chan JY, Dabrowski-Amaral CE, Delvecchio AM, Dillon SB;
 XX Leary JJ;

DR WPI; 1998-286847/25.
 XX

DR N-PSDB; AAV62159.
 XX

PT Herpes simplex virus type-2 sequences - useful in, e.g. prevention and
 XX treatment of infection or inducing immunological response in mammal.

PS Claim 10; Page 92; 748pp; English.
 XX

CC This sequence represents a Herpes simplex virus type-2 (HSV-2) protein
 CC sequence of the invention. This sequence was isolated from a HSV-2 strain
 CC SB5 (deposited as ATCC VR-2546) DNA fragment designated Contig ID 15.
 CC Based on homology, this sequence is a UL26 protein. The proteins can be
 CC used for the treatment or prevention of disease, to induce an
 CC immunological response in a mammal or to identify inhibitors, activators
 CC or novel antivirals. Antagonists of the proteins can be used to inhibit a
 CC viral polypeptide. The DNA sequence or a vector containing it can also be
 CC used to induce an immunological response in a mammal

SQ Sequence 642 AA;

Query Match 99.2%; Score 3359.5; DB 2; Length 642;
 Best Local Similarity 99.4%; Pred. No. 1.7e-207;

Matches 634; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 1 MASAEWRERLEAPLPDRAVPIYVAGFLALYDSGDPGELALDPTVTRAAALPPENPLPINVD 60
DB 5 MASAEWRERLEAPLPDRAVPIYVAGFLALYDSGDPGELALDPTVTRAAALPPENPLPINVD 64
QY 61 HRARCEVGRVLA VNDPRGPFVGLIACVQLSERVLETAASAAI FERRGPALSRERLLYL 120
DB 65 HRARCEVGRVLA VNDPRGPFVGLIACVQLSERVLETAASAAI FERRGPALSRERLLYL 124
QY 121 ITNYLPSVLSLTKRGDEVPDRTILFAHVALCAIGRRLGTIVTYDTSDDAAIAPFRHLDP 180
DB 125 ITNYLPSVLSLTKRGDEVPDRTILFAHVALCAIGRRLGTIVTYDTSDDAAIAPFRHLDP 184
QY 181 ATREGVRRERAAEALAGRTWAPGVEALHTLLSTAVNNMLDRWSLVAERRRQAGTA 240
DB 185 ATREGVRRERAAEALAGRTWAPGVEALHTLLSTAVNNMLDRWSLVAERRRQAGTA 244
QY 241 GHTYLOASEKFKIWGAESAPAPERGYKTGAPGAMDTSPAASVPAQVAVRARAQVA -SSSS 299
DB 245 GHTYLOASEKFKIWGAESAPAPERGYKTGAPGAMDTSPAASVPAQVAVRARAQVA-SSSS 304
QY 300 SSSFPAPADMNPFVSASGAPAPPPPGDGSYLWIPASHYNQLVTGQAPRHPPLTACGLPAA 359
DB 305 SSSFPAPADMNPFVSASGAPAPPPPGDGSYLWIPASHYNQLVTGQAPRHPPLTACGLPAA 364
QY 360 GTVAYGHGPGAGSPHYPPPPPAHPYPCMLPAGSPSLEAQIAALVGAIAADRQAGGLPAAAG 419
DB 365 GTVAYGHGPGAGSPHYPPPPPAHPYPCMLPAGSPSLEAQIAALVGAIAADRQAGGLPAAAG 424
QY 420 DHGIRGSAKRRRHEVEQPEYDCGRDEPDRDPFYYPGEARPEPRVDSRRAARQASGPHE 479
DB 425 DHGIRGSAKRRRHEVEQPEYDCGRDEPDRDPFYYPGEARPEPRVDSRRAARQASGPHE 484
QY 480 ITALVGAVTSLQOELAHMRARTHAPYGPYPVPGYHHPHADTETPAQPPRYPAKAVYLP 539
DB 485 ITALVGAVTSLQOELAHMRARTHAPYGPYPVPGYHHPHADTETPAQPPRYPAEAVYLP 544
QY 540 PHIAPPGPPLSGAVPPSPYPVAVTGPAPPLHQPSPAHAHPPPPPPGTPPPPAASLPQP 599
DB 545 PHIAPPGPPLSGAVPPSPYPVAVTGPAPPLHQPSPAHAHPPPPPPGTPPPPAASLPQP 604
QY 600 EAPGAEGALVNASSAAHVNDTARAADLFVSMGMSR 637
DB 605 EAPGAEGALVNASSAAHVNDTARAADLFVSMGMSR 642

RESULT 5
AAR64767
ID AAR64767 standard; protein; 636 AA.

AC AAR64767;
DT 25-MAR-2003 (revised)
DT 19-JUL-1995 (first entry)

DE HSV-2 protease, ICP35.
KW Protease; HSV-2; protease-inhibitor; virucide; ICP35.
XX Herpes simplex virus type 2.

OS

XX WO9429456-A2.
XX 22-DEC-1994.
XX 25-MAY-1994; 94WO-US005920.

XX 08-JUN-1993; 93US-00073819.
XX 23-MAY-1994; 94US-00245390.

XX (ABBO) ABBOTT LAB.

PI Steffy KR, Kati WM, Katz L, Mgonigal TP, Sarthy AV, Schoen SE;
XX WPI; 1995-036483/05.
DR N-PSDB; AAR64767.
XX New Herpes Simplex Virus type 2 protease - used in screening methods for
PT identifying potential herpes viral protease inhibitor cpds.
XX Claim 4; Fig 1A-1C; 56pp; English.
XX HSV-2 DNA (given in AAR64767) was isolated from Vero cells infected with
CC HSV-2 strain Ci. The translated amino acid sequence is given in AAR64767.
CC The DNA was digested with BamHI, sepd. by gel electrophoresis and
CC transferred to nitrocellulose. A nick-translated probe deriv. from HSV-1
CC protease was used for hybridization to identify DNA (AAR64767) encoding
CC HSV-2 protease (AAR64768) and the protease substrate ICP35 (AAR64769).
CC (Updated on 25-MAR-2003 to correct PN field.)
XX Sequence 636 AA;

Query Match 99.0%; Score 3351.5; DB 2; Length 636;
Best Local Similarity 99.4%; Pred. No. 5.6e-207;
Matches 633; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 1 MASAEWRERLEAPLPDRAVPIYVAGFLALYDSGDPGELALDPTVTRAAALPPENPLPINVD 60
DB 1 MASAEWRERLEAPLPDRAVPIYVAGFLALYDSGDPGELALDPTVTRAAALPPENPLPINVD 60
QY 61 HRARCEVGRVLA VNDPRGPFVGLIACVQLSERVLETAASAAI FERRGPALSRERLLYL 120
DB 61 HRARCEVGRVLA VNDPRGPFVGLIACVQLSERVLETAASAAI FERRGPALSRERLLYL 120
QY 121 ITNYLPSVLSLTKRGDEVPDRTILFAHVALCAIGRRLGTIVTYDTSDDAAIAPFRHLDP 180
DB 121 ITNYLPSVLSLTKRGDEVPDRTILFAHVALCAIGRRLGTIVTYDTSDDAAIAPFRHLDP 180
QY 181 ATREGVRRERAAEALAGRTWAPGVEALHTLLSTAVNNMLDRWSLVAERRRQAGTA 240
DB 181 ATREGVRRERAAEALAGRTWAPGVEALHTLLSTAVNNMLDRWSLVAERRRQAGTA 240
QY 241 GHTYLOASEKFKIWGAESAPAPERGYKTGAPGAMDTSPAASVPAQVAVRARAQVA-SSSS 300
DB 241 GHTYLOASEKFKIWGAESAPAPERGYKTGAPGAMDTSPAASVPAQVAVRARAQVA-SSSS 299
QY 301 SSSFPAPADMNPFVSASGAPAPPPPGDGSYLWIPASHYNQLVTGQAPRHPPLTACGLPAA 360
DB 300 SSSFPAPADMNPFVSASGAPAPPPPGDGSYLWIPASHYNQLVTGQAPRHPPLTACGLPAA 359
QY 361 TVAYGHGPGAGSPHYPPPPPAHPYPCMLPAGSPSLEAQIAALVGAIAADRQAGGLPAAAG 420
DB 360 TVAYGHGPGAGSPHYPPPPPAHPYPCMLPAGSPSLEAQIAALVGAIAADRQAGGLPAAAG 419
QY 421 HGIRGSAKRRRHEVEQPEYDCGRDEPDRDPFYYPGEARPEPRVDSRRAARQASGPHE 480
DB 420 HGIRGSAKRRRHEVEQPEYDCGRDEPDRDPFYYPGEARPEPRVDSRRAARQASGPHE 479
QY 481 TALVGAVTSLQOELAHMRARTHAPYGPYPVPGYHHPHADTETPAQPPRYPAKAVYLP 540
DB 480 TALVGAVTSLQOELAHMRARTHAPYGPYPVPGYHHPHADTETPAQPPRYPAEAVYLP 539
QY 541 HIAPPGPPLSGAVPPSPYPVAVTGPAPPLHQPSPAHAHPPPPPPGTPPPPAASLPQP 600
DB 540 HIAPPGPPLSGAVPPSPYPVAVTGPAPPLHQPSPAHAHPPPPPPGTPPPPAASLPQP 599
QY 601 AFGAEGALVNASSAAHVNDTARAADLFVSMGMSR 637
DB 600 AFGAEGALVNASSAAHVNDTARAADLFVSMGMSR 636

RESULT 6
AAW72192
ID AAW72192 standard; protein; 657 AA.
XX

AC AAW72192;
 XX 13-JAN-1999 (first entry)
 XX HSV-2 strain SB5 Contig ID 15 ORF#27 protein.
 XX HSV-2 strain SB5; immunological response induction; therapy;
 KW antiviral identification; viral protein inhibitor.
 XX Herpes simplex virus 2.
 XX WO9820016-A1.
 XX 14-MAY-1998.
 XX 31-OCT-1997; 97WO-US020016.
 XX 04-NOV-1996; 96US-0030279P.
 PR 09-JUN-1997; 97US-0049018P.
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 XX Esser KM, Chan JY, Dabrowski-Amarel CE, Delvecchio AM, Dillon SB;
 PI Leary JJ;
 XX WPI; 1998-286847/25.
 DR N-PSDB; AAV62176.
 XX Herpes simplex virus type-2 sequences - useful in, e.g. prevention and
 PT treatment of infection or inducing immunological response in mammal.
 XX Claim 10; Page 120; 748pp; English.
 XX This sequence represents a Herpes simplex virus type-2 (HSV-2) protein
 CC sequence of the invention. This sequence was isolated from a HSV-2 strain
 CC SB5 (deposited as ARCC VR-2546) DNA fragment designated Contig ID 15.
 CC Based on homology, this sequence is a capsid protein P40. The proteins
 CC can be used for the treatment or prevention of disease, to induce an
 CC immunological response in a mammal or to identify inhibitors, activators
 CC or novel antivirals. Antagonists of the proteins can be used to inhibit a
 CC viral polypeptide. The DNA sequence or a vector containing it can also be
 CC used to induce an immunological response in a mammal
 XX
 XX Sequence 657 AA;

Query Match 96.5%; Score 3267.5; DB 2; Length 657;
 Best Local Similarity 99.4%; Pred. No. 1.4e-201;
 Matches 615; Conservative 1; Mismatches 2; Indels 1; Gaps 1;
 QY 1 MASAEMRERLEAPLDRAPVPIYVAGFLALYDSGDPGELALDPTVRAALPPENPLPINVD 60
 DB 1 MASAEMRERLEAPLDRAPVPIYVAGFLALYDSGDPGELALDPTVRAALPPENPLPINVD 60
 QY 61 HRARCEVGRVLAVNDPRGPFVGLIACVOLERVLETAASAAIFERRGPALSREERLLYL 120
 DB 61 HRARCEVGRVLAVNDPRGPFVGLIACVOLERVLETAASAAIFERRGPALSREERLLYL 120
 QY 121 ITNVLPSVLSLTKRGDEVPDRTLFAHVALCAIGRLGTIVTYDTSLDAAIAPFRHLDP 180
 DB 121 ITNVLPSVLSLTKRGDEVPDRTLFAHVALCAIGRLGTIVTYDTSLDAAIAPFRHLDP 180
 QY 181 ATREGVRRERAAEALAGTWTAPGVLEALHTLLSTAVNNMMLDRWSLVAERRRONGIA 240
 DB 181 ATREGVRRERAAEALAGTWTAPGVLEALHTLLSTAVNNMMLDRWSLVAERRRONGIA 240
 QY 241 GHTYLQASEFKIWGAEAPAPRGYKTAGPAMDTSPAASVPAPQVAVRQVA-SSSS 299
 DB 241 GHTYLQASEFKIWGAEAPAPRGYKTAGPAMDTSPAASVPAPQVAVRQVASSSS 300
 QY 300 SSSFPAPADNPNVSASGAPAPPDGGSYLWIPASHYNQLVTGQSAPRHPPLTACGLPAA 359
 DB 301 SSSFPAPADNPNVSASGAPAPPDGGSYLWIPAFHYNQLVTGQSAPRHPPLTACGLPAA 360

QY 360 GTVAYGHPGAGSPSPHYPPPPPAHPYPGMLFAGPSPLEAQIAALVGAIADRAQAGLPAAAG 419
 DB 361 GTVAYGHPGAGSPSPHYPPPPPAHPYPGMLFAGPSPLEAQIAALVGAIADRAQAGLPAAAG 420
 QY 420 DHGIRGSAKRERHEVEQPEYDCGRDPRDPFPYFGEARPEPRPVDSSRAARQASGPHET 479
 DB 421 DHGIRGSAKRERHEVEQPEYDCGRDPRDPFPYFGEARPEPRPVDSSRAARQASGPHET 480
 QY 480 ITALVGAVTSLQQLAHMRARTHAPYCGYPVPGVGYHHPHADTETPAQPPRYPAKAVYLP 539
 DB 481 ITALVGAVTSLQQLAHMRARTHAPYCGYPVPGVGYHHPHADTETPAQPPRYPAKAVYLP 540
 QY 540 PHIAAPPGLPLSGAVPPSPVPPVAVTGPAPPLHQPSPAHAPHPPPPPPPTPPPAASLPQP 599
 DB 541 PHIAAPPGLPLSGAVPPSPVPPVAVTGPAPPLHQPSPAHAPHPPPPPPPTPPPAASLPQP 600
 QY 600 EAPGAENGALVNASSAAHV 618
 DB 601 EAPGAENGALVNASSAAHV 619
 RESULT 7
 AAR71031
 ID AAR71031 standard; protein; 635 AA.
 XX AAR71031;
 AC 25-MAR-2003 (revised)
 DT 02-OCT-1995 (first entry)
 XX HSV-1 UL26 gene product.
 DE Herpes simplex virus; HSV-1 protease; capsid.
 KW Herpes simplex virus type 1.
 OS Herpes simplex virus type 1.
 XX WO9506055-A1.
 PN 02-MAR-1995.
 PD 19-AUG-1994; 94WO-US009303.
 XX 20-AUG-1993; 93US-00110522.
 PR 23-JUN-1994; 94US-00264537.
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 PA Dilella AG, Debouck CM;
 XX WPI; 1995-106803/14.
 DR N-PSDB; AAQ84678.
 XX New herpes simplex virus (HSV)-2 protease and capsid protein - used to
 PT develop prods. for use in the diagnosis and treatment of HSV-2
 PT infections.
 XX Disclosure; Page 35; 51pp; English.
 PS The sequence is that of a herpes simplex virus type 1 protease and
 CC capsid. The protein can be used in the diagnosis and treatment of HSV-1
 CC infections. See also R717017-30. (Updated on 25-MAR-2003 to correct PN
 CC field.)
 XX Sequence 635 AA;
 SQ Query Match 68.9%; Score 2333; DB 2; Length 635;
 Best Local Similarity 74.0%; Pred. No. 1.4e-141;
 Matches 477; Conservative 30; Mismatches 120; Indels 18; Gaps 7;
 QY 1 MASAEMRERLEAPLDRAPVPIYVAGFLALYDSGDPGELALDPTVRAALPPENPLPINVD 60
 DB 1 MASAEMRERLEAPLDRAPVPIYVAGFLALYDSGDPGELALDPTVRAALPPENPLPINVD 60

QY 535 VYLPPIHAPGPPPLSGA--VPPSYPPVAVTPGAPPLHQPSPAHAPPPPPGPTPPP 592
 DB 534 VYRPPPHSAFYGGPPGASHAPTPPYAPACPGGPPP---PCPSTQTRAPLTPTEPAFP 590
 QY 593 AASLPQPEAGAEAGALVNASSAAHVNDTARAADLFVSQMMGSR 637
 DB 591 AATGSQPEASNAEAGALVNASSAAHVNDTARAADLFVSQMMGAR 635

RESULT 9

AAR28634
 ID AAR28634 standard; protein; 635 AA.
 AC AAR28634;
 DT 25-MAR-2003 (revised)
 DT 24-MAR-1993 (first entry)
 DE UL26 protease.
 KW UL26; open reading frame; ORF; protease; herpes simplex virus; HSV;
 KW capsid protein; ICP35; ICP35 c; ICP35 d; ICP35 e; ICP35 f;
 KW functional domain.
 OS Herpes simplex virus type 1.
 XX EP514830-A2.
 XX 25-NOV-1992.
 XX 19-MAY-1992; 92EP-00108420.
 XX 24-MAY-1991; 91US-00705814.
 PR 07-FEB-1992; 92US-00832855.
 XX (ARCH-) ARCH DEV CORP.
 PA Roizman B, Liu F;
 PI WPI; 1992-391444/48.
 DR N-PSDB; AAQ30736.

XX New herpes serine proteases and corresp. nucleic acid sequences - for
 PT detection, prevention and treatment of infection caused by HSV, Epstein-
 PT Barr, Varicella-Zoster and CMV cytomegalo-virus.
 XX Disclosure; Fig 1B; 66pp; English.
 XX This sequence is encoded by the UL26 open reading frame (ORF) and it is a
 CC protease. The UL26 ORF is contained within a section of the herpes
 CC simplex virus (HSV) type 1 genome for the family of HSV capsid proteins
 CC designated ICP35. The UL26 transcription initiation site is used as the
 CC start point for measurements of distance in this section of the genome.
 CC The UL26 protease and ICP35 precursor proteins, ICP35 c and d, may be
 CC cleaved by the UL26 protease. Cleavage of ICP35 c and d gives ICP35 e and
 CC f respectively. Cleavage specificity of the UL26 protease may be altered
 CC by mutations within the sequence. These mutations may be insertions,
 CC deletions or substitutions which affect the functional domains of the
 CC protease. See also AAR28635-52. (Updated on 25-MAR-2003 to correct PN
 CC field.)
 XX SQ Sequence 635 AA;

Query Match 68.8%; Score 2329; DB 2; Length 635;
 Best Local Similarity 73.8%; Pred. No. 2.5e-141;
 Matches 476; Conservative 30; Mismatches 121; Indels 18; Gaps 7;
 QY 1 MASAEMRERLEAPDRAVPIYVAGFLYDSDGELALDPTVRAALPPENPLPNTVD 60
 DB 1 MAADAPGRMEEPDRAVPIYVAGFLYDSDGELALDPTVRAALPPDPLPNTVD 60
 QY 61 HRACEVGRVLAVVNDPRGPFVGLIACVQLERVLETAASAAIFERRGPALSREERLLYL 120

DB 61 HRACEVGRVLAVVNDPRGPFVGLIACVQLERVLETAASAAIFERRGPALSREERLLYL 120
 QY 121 ITNLYSVSLSTKRGDEVPPDRTLFAHVALCAIGRLGTIVYDTSLDAAIAFPRLDP 180
 DB 121 ITNLYSVSLATKRLGGEAHPDRTLFAHVALCAIGRLGTIVYDGLDAAIAFPRLSP 180
 QY 181 ATREGVRRRAAEALALAGRTWAPGVEALHTLLSTAVNNMLDRMSLVAERRRQAGIA 240
 DB 181 ASREGARRLAEEALALSGRTWAPGVEALHTLLSTAVNNMLDRMSLVAERRRQAGIA 240
 QY 241 GHTYLOASEKPKIWAESAPAPERGYKTGAPGAMDTSPAASVPAQ-----VAVRAROVAS 296
 DB 241 GHTYLOASEKFKWGAEPVSAPARGYKNGAPESTDIPPGSIAAAPQGDRCPIVRQRVAL 300
 QY 297 SSSSSFFPAPADNPNVSASGAPAPPDGGSGSYLWIPASHYNQLVGTGSAAPRHPLTACGL 356
 DB 301 S-----PVLPPNVPVTSPTAPAPDGGSGSYLWIPASHYNQLVAGHAAPQDPHSAFGP 354
 QY 357 P-AAGTVAYGHPGAGPSPHYPPPPHAPYPCGMLFAGSPLEAQIAALVGAIAAARQAGLP 415
 DB 355 PAAAGSVAYGPHGAGLSQHYPPHVAHQYPCGLFSGSPLEAQIAALVGAIAAARQAGQP 414
 QY 416 AAAGDHGIRGSAKRRHEVEQPEYDCGRDEPDRDPFYPCGEARPEPRPVDSRAARQASG 475
 DB 415 -AAGDPGVRGSGKRRRYEAGPSESYCDQDEPDADYPYPCGEARGAPRGVDSRAARHSPG 473
 QY 476 PHETITLVGAVTSLQOELAHMEARTHAPYGPVPPYVPHHHPADTE-TPAQPPRYPAKA 534
 DB 474 TNETITLVGAVTSLQOELAHMEARTHAPYGPVPPYVPHHHPADTE-TPAQPPRYPAKA 533
 QY 535 VYLPPIHAPGPPPLS--GAVPPPSYPVAVTPGAPPLHQPSPAHAPPPPPGPTPPP 592
 DB 534 VYRPPPHSAFYGGPPGASHAPTPPYAPACPGGPPP---PCPSTQTRAPLTPTEPAFP 590
 QY 593 AASLPQPEAGAEAGALVNASSAAHVNDTARAADLFVSQMMGSR 637
 DB 591 AATGSQPEASNAEAGALVNASSAAHVNDTARAADLFVSQMMGAR 635

RESULT 10

AAR28651
 ID AAR28651 standard; protein; 635 AA.
 AC AAR28651;
 DT 25-MAR-2003 (revised)
 DT 24-MAR-1993 (first entry)
 DE UL26 protease substitution mutant KK.
 KW UL26; open reading frame; ORF; protease; herpes simplex virus; HSV;
 KW capsid protein; ICP35; ICP35 c; ICP35 d; ICP35 e; ICP35 f;
 KW functional domain; insertion; deletion; substitution.
 OS Herpes simplex virus type 1.
 FH Key Location/Qualifiers
 FT Misc-difference 215 /label= S215A
 FT EP514830-A2.
 XX 25-NOV-1992.
 XX 19-MAY-1992; 92EP-00108420.
 XX 24-MAY-1991; 91US-00705814.
 PR 07-FEB-1992; 92US-00832855.
 XX (ARCH-) ARCH DEV CORP.
 PA Roizman B, Liu F;
 PI

Db 181 ASREGARRLAABAEALSGRTWARGVEALHTLLSTAVNNMLDRWSLVAERRRQAGIA 240
 Qy 241 GHTYLOASEFKIWAESAEPAPERYKGTGAPGAMDTSPAASVAPAQ-----VAVRARQVAS 296
 Db 241 GHTYLOASEFKIWAESAEPAPERYKGTGAPGAMDTSPAASVAPAQ-----VAVRARQVAS 300
 Qy 297 SSSSSFPADMPVSAAGAPAPPPGDSYLMIPASHYNQLVTGQAPRHPPLTACGL 356
 Db 301 S-----PVLPPMNPVPTSGTPAPPGDSYLMIPASHYNQLVTGQAPRHPPLTACGL 354
 Qy 357 P-AAGTVAYGHPGAGSPHPPPPAHPVPGMLFAGSPLEAQIAALVGAIAADRQAGGLP 415
 Db 355 PAAAGSVAYGHPGAGLSQHTPPHVAHQYPGVLFSPGLEAQIAALVGAIAADRQAGQP 414
 Qy 416 AAAGDHGIRGSAKRRRHEVEQPEYDCRDEPDRTLPYYPGEARPEPRVDSRRAARQASG 475
 Db 415 -AAGDPGVRGSKRRRYEAGPSESYCDQDEPDADYPYYPGEARGAPRGVDSRRAARHSPG 473
 Qy 476 PHETITALMGAVTSLOQELAHMRARTAPYGPYPVPGYHHPHADTTP--AQPPRYPAK 533
 Db 474 TNETITALMGAVTSLOQELAHMRARTAPYGMVTPVAYRXPQVGEPEPTTTHPALCPPE 533
 Qy 534 AVYLPPIHAPPGPPLS--GAVPPSPYPVAVTPGAPPLHQPSPAHAHPPPPPGPTTP 591
 Db 534 AVYRPPHSAFYGPPQGPSPHAPTPPYAPACPPGPPPP---PCPSTQTRAPLPTPEPAFP 590
 Qy 592 PAASLOPEAPGAEAGALVNASSAAHVNVDTRAADLTVSOMMGRS 637
 Db 591 PAATGSQPEASNAEAGALVNASSAAHVDVDTARAADLTVSOMMGR 636

RESULT 12

AAR28647
 ID AAR28647 standard; protein; 635 AA.

AC AAR28647;
 XX

XX 25-MAR-2003 (revised)

DT 24-MAR-1993 (first entry)

DE UL26 protease substitution mutant GG.

XX UL26; open reading frame; ORF; protease; herpes simplex virus; HSV;
 KW capsid protein; ICP35; ICP35 c; ICP35 d; ICP35 e; ICP35 f;
 KW functional domain; insertion; deletion; substitution.

OS Herpes simplex virus type 1.

XX Key Location/Qualifiers
 FH Misc-difference 7.9
 FT /note= "Substitution GDR > SRT"

XX EF514830-A2.

XX 25-NOV-1992.

XX 19-MAY-1992; 92EP-00108420.

XX 24-MAY-1991; 91US-00705814.

XX 07-FEB-1992; 92US-00832855.

XX (ARCH-) ARCH DEV CORP.

XX Roizman B, Liu F;

XX WPI; 1992-39144/48.

XX New herpes serine proteases and corresp. nucleic acid sequences - for

XX detection, prevention and treatment of infection caused by HSV, Epstein-

XX Barr, Varicella-Zoster and CMV cytomegalo-virus.

XX Disclosure; Table 1; 66pp; English.

XX PS

XX

CC The sequences given in AAR28635-52 are mutants of the protease encoded by
 CC the UL26 open reading frame (ORF). The UL26 ORF is contained within a
 CC section of the herpes simplex virus (HSV) type 1 genome for the family of
 CC HSV capsid proteins designated ICP35. The UL26 transposition initiation
 CC site is used as the start point for measurements of distance in this
 CC section of the genome. The UL26 protease and ICP35 precursor proteins,
 CC ICP35 c and d, may be cleaved by the UL26 protease. Cleavage of ICP35 c
 CC and d gives ICP35 e and f respectively. Cleavage specificity of the UL26
 CC protease is altered by the inclusion of these mutations. These mutations
 CC include insertions, deletions or substitutions which affect the functional
 CC domains of the protease. (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 635 AA;

Query Match 68.6%; Score 2322; DB 2; Length 635;

Best Local Similarity 73.6%; Pred. No. 7e-141;

Matches 475; Conservative 29; Mismatches 123; Indels 18; Gaps 7;

Qy 1 MASAEMRERLEAPLDRAPVIYVAGFLALYDSGDPGLALDDPTVRAALPENPLPIND 60
 Db 1 MAADAPSRRTMEELPDRAPVIYVAGFLALYDSGDPGLALDDPTVRAALPENPLPIND 60
 Qy 61 HRARCEVGRVLAVNDPRGPFVGLIACVOLERVLETAASAAIFEREGPALSRERLLYL 120
 Db 61 HRACGEVGRVLAVNDPRGPFVGLIACVOLERVLETAASAAIFEREGPALSRERLLYL 120
 Qy 121 ITNYPVSVLSLTKRRGDEVDPDRTLFAHVALCAIGRLGTIVTYDTSIDAIAAPFRHLDP 180
 Db 121 ITNYPVSVLSLTKRRGDEVDPDRTLFAHVALCAIGRLGTIVTYDTSIDAIAAPFRHLDP 180
 Qy 181 ATREGVREAAEAELALAGRTWAPGVEALHTLLSTAVNNMLDRWSLVAERRRQAGIA 240
 Db 181 ASREGARRLAABAEALSGRTWAPGVEALHTLLSTAVNNMLDRWSLVAERRRQAGIA 240
 Qy 241 GHTYLOASEFKIWAESAEPAPERYKGTGAPGAMDTSPAASVAPAQ-----VAVRARQVAS 296
 Db 241 GHTYLOASEFKIWAESAEPAPERYKGTGAPGAMDTSPAASVAPAQ-----VAVRARQVAS 300
 Qy 297 SSSSSFPADMPVSAAGAPAPPPGDSYLMIPASHYNQLVTGQAPRHPPLTACGL 356
 Db 301 S-----PVLPPMNPVPTSGTPAPPGDSYLMIPASHYNQLVTGQAPRHPPLTACGL 354
 Qy 357 P-AAGTVAYGHPGAGSPHPPPPAHPVPGMLFAGSPLEAQIAALVGAIAADRQAGGLP 415
 Db 355 PAAAGSVAYGHPGAGLSQHTPPHVAHQYPGVLFSPGLEAQIAALVGAIAADRQAGQP 414
 Qy 416 AAAGDHGIRGSAKRRRHEVEQPEYDCRDEPDRTLPYYPGEARPEPRVDSRRAARQASG 475
 Db 415 -AAGDPGVRGSKRRRYEAGPSESYCDQDEPDADYPYYPGEARGAPRGVDSRRAARHSPG 473
 Qy 476 PHETITALMGAVTSLOQELAHMRARTAPYGPYPVPGYHHPHADTTP--TPAQRPPRYPAKA 534
 Db 474 TNETITALMGAVTSLOQELAHMRARTAPYGMVTPVAYRXPQVGEPEPTTTHPALCPPEA 533
 Qy 535 VYLPPIHAPPGPPLS--GAVPPSPYPVAVTPGAPPLHQPSPAHAHPPPPPGPTTPP 592
 Db 534 VYRPPHSAFYGPPQGPSPHAPTPPYAPACPPGPPPP---PCPSTQTRAPLPTPEPAFP 590
 Qy 593 AASLOPEAPGAEAGALVNASSAAHVNVDTRAADLTVSOMMGRS 637
 Db 591 AATGSQPEASNAEAGALVNASSAAHVDVDTARAADLTVSOMMGR 635

RESULT 13

AAR28652
 ID AAR28652 standard; protein; 635 AA.

XX AAR28652;
 AC

XX 25-MAR-2003 (revised)

DT 24-MAR-1993 (first entry)

XX UL26 protease substitution mutant LL.

DE

Query Match	68.5%; Score 2319; DB 2; Length 626;
Best Local Similarity	74.4%; Pred. No. 1.1e-140;
Matches	473; Conservative 28; Mismatches 117; Indels 18; Gaps 7;
QY	10 LEAPLPDRAVPIYVAGFLALYDSDGDELALDPTVRAALPPENPLINVDHRCCEVGR 69
DB	1 MEBPLDRAVPIYVAGFLALYDSDGDELALDPTVRAALPPDNPLINVDHRCCEVGR 60
QY	70 VLA VNDPRGFFVGLIACVQLERVLTAASAAIFERRGPALGREERLLYLITNYPVS 129
DB	61 VLA VDDPRGFFVGLIACVQLERVLTAASAAIFERRGPALGREERLLYLITNYPVS 120
QY	130 LSTKRGEVPPDRTLFAHVALCAIGRLGTIVTYDTSLDAAIAPFRHLDPATREGVRE 189
DB	121 LATKRLGGEAHPDRTLFAHVALCAIGRLGTIVTYDTSGLDAAIAPFRHLSPASREGARRL 180
QY	190 AAEALALAGRTWAPGVEALTHTLSTAVNNMLDRWSLVAERROAGIAGHTYLOASE 249
DB	181 AAEALALSGRTWAPGVEALTHTLSTAVNNMLDRWSLVAERROAGIAGHTYLOASE 240
QY	250 KFKIWGAESAPAPRGYKTCAPGAMDTSPAASVPAPQ---VAVRAQVASSSSSSPPA 305
DB	241 KFKWGAEPVSAPARGYKNGAPESTDIPGSIAPAPQDRCPIVRQGVALS-----PV 294
QY	306 PADMNVPVASGAPAPPDPGSGSYLWIPASHYNQLVTQSPAPRHPLLTACGLP-AAGTVAY 364
DB	295 LPPMNPVPTSGTPAPAPPDPGSGSYLWIPASHYNQLVAGHAAPQPQPHSAFGFPAAAGSVAY 354
QY	365 GHGAGSPHYPPPPAHYPGMLFAGSPLEAQIAALVGAIAADRQAGGLPAAAGDHGIR 424
DB	355 GPFGAGLSQHYPPHVAHQYGVLFSGSPLEAQIAALVGAIAADRQAGGQP-AAGDPGVR 413
QY	425 GSARRRHEVEQPEYDCGRDEPRDPYYPGEARPEPRPVDSRAARQASGHETITIALV 484
DB	414 GSKRRRYEAGPSYSCDQDEPADPIYPYEGEARGPRGVDSRAARKHSGTNETITALM 473
QY	485 GAVTSLQOELAHMRARTHAPYGPYPVPYVGYPHHDPADTE-TPAQPPRYPAKAVYLPHPHIA 543
DB	474 GAVTSLQOELAHMRARTSAFYGMVTPVAHYRPQVGEPEPTTTHPALCPPEAVYRPPHSA 533
QY	544 PPGPPLS--GAVPPSPYPPVAVTPGAPPLHQSPAHAPPPPPPGTPPPPAASLPQPEA 601
DB	534 PYGPPQGPSPHAPTTPYAPAACPGPPPPP---PCPSTQTRAPLTSPAPFPAPATGSOPEA 590
QY	602 PGAEAGALVNASSAAHVNDTARAADLFVSQMMGSR 637
DB	591 SNAEAGALVNASSAAHVNDTARAADLFVSQMMGAR 626
RESULT 15	
AA28650	
ID	AA28650 standard; protein; 635 AA.
XX	
AC	AA28650;
XX	
DT	25-MAR-2003 (revised)
DT	24-MAR-1993 (first entry)
XX	
DE	UL26 protease substitution mutant JJ.
XX	
KW	UL26; open reading frame; ORF; protease; herpes simplex virus; HSV;
KW	capsid protein; ICP35; ICP35 c; ICP35 d; ICP35 e; ICP35 f;
KW	functional domain; insertion; deletion; substitution.
XX	
OS	Herpes simplex virus type 1.
XX	
Key	Location/Qualifiers
FT	Misc-difference 148
FT	/label= H148A
XX	
PN	EP514830-A2.
XX	

PD	25-NOV-1992.	
XX		
PF	19-MAY-1992; 92EP-00108420.	
XX		
PR	24-MAY-1991; 91US-00705814.	
PR	07-FEB-1992; 92US-00832855.	
XX		
PA	(ARCH-) ARCH DEV CORP.	
XX		
PI	Roizman B, Liu F;	
XX		
DR	WPI; 1992-391444/48.	
XX		
PT	New herpes serine proteases and corresp. nucleic acid sequences - for	
PT	detection, prevention and treatment of infection caused by HSV, Epstein-	
PT	Barr, Varicella-Zoster and CMV cytomegalo-virus.	
XX		
PS	Disclosure; Table 1; 66pp; English.	
XX		
CC	The sequences given in AAR28635-52 are mutants of the protease encoded by	
CC	the UL26 open reading frame (ORF). The UL26 ORF is contained within a	
CC	section of the herpes simplex virus (HSV) type 1 genome for the family of	
CC	HSV capsid proteins designated ICP35. The UL26 transcription initiation	
CC	site is used as the start point for measurements of distance in this	
CC	section of the genome. The UL26 protease and ICP35 precursor proteins,	
CC	ICP35 c and d, may be cleaved by the UL26 protease. Cleavage of ICP35 c	
CC	and d gives ICP35 e and f respectively. Cleavage specificity of the UL26	
CC	protease is altered by the inclusion of these mutations. These mutations	
CC	include insertions, deletions or substitutions which affect the functional	
CC	domains of the protease. (Updated on 25-MAR-2003 to correct PN field.)	
XX		
SQ	Sequence 635 AA;	
Query Match		68.5%; Score 2319; DB 2; Length 635;
Best Local Similarity		73.6%; Pred. No. 1.1e-140;
Matches		475; Conservative 30; Mismatches 122; Indels 18; Gaps 7;
QY	1	MASAEERLERLEAPLPDRAVPIYVAGFLALYDSDGDELALDPTVRAALPPENPLINVD 60
DB	1	MAADAPGDRMEELPLDRAVPIYVAGFLALYDSDGDELALDPTVRAALPPDNPLINVD 60
QY	61	HRARCEVGRVLA VNDPRGFFVGLIACVQLERVLTAASAAIFERRGPALGREERLLYL 120
DB	61	HRAGCEVGRVLA VNDPRGFFVGLIACVQLERVLTAASAAIFERRGPALGREERLLYL 120
QY	121	ITNYPVSLSLTKRCGEVPPDRTLFAHVALCAIGRLGTIVTYDTSLDAAIAPFRHLDP 180
DB	121	ITNYPVSLSLTKRLGGEAHPDRTLFAAVALCAIGRLGTIVTYDTSGLDAAIAPFRHLSP 180
QY	181	ATREGVRRREAAAEALALAGRTWAPGVEALTHTLSTAVNNMLDRWSLVAERROAGIA 240
DB	181	ASREGARRLLAAEALALSGRTWAPGVEALTHTLSTAVNNMLDRWSLVAERROAGIA 240
QY	241	GHTYLOASEKFKTWGAESAPAPRGYKTCAPGAMDTSPAASVPAPQ---VAVRAQVAS 296
DB	241	GHTYLOASEKFKWGAEPVSAPARGYKNGAPESTDIPGSIAPAPQDRCPIVRQGVAL 300
QY	297	SSSSSFPAPADMNVPVSASGAPAPPPGSGSYLWIPASHYNQLVTQSPAPRHPLTACGL 356
DB	301	S-----PVLPPMNPVPTSGTPAPAPPDPGSGSYLWIPASHYNQLVAGHAAPQPQPHSAFGF 354
QY	357	P-AAGTVAYVGHPCGAGSPHYPPPPAHYPGMLFAGSPLEAQIAALVGAIAADRQAGGLP 415
DB	355	PAAGSVAVGPHGAGLSQHYPPHVAHQYGVLFSGSPLEAQIAALVGAIAADRQAGGQP 414
QY	416	AAAGDHGIRGSAKRRRHEVEQPEYDCGRDEPRDPYYPGEARPEPRPVDSRAARQASG 475
DB	415	-AAGDPGVGSGKRRRYEAGPSYSCDQDEPADPIYPYEGEARGPRGVDSRAARHSPG 473
QY	476	PHETITIALVCANTSLQOELAHMRARTHAPYGPYPVPYVGYPHHDPADTE-TPAQPPRYPAKA 534
DB	474	TNETITIALGAVTSLQOELAHMRARTSAFYGMVTPVAHYRPQVGEPEPTTTHPALCPPEA 533

